numero 215
luglio 1986
anno XIX
volume XXXVII

LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

La morte improvvisa cardiaca

Il cuore può smettere di pompare il sangue senza alcun segno premonitore: il soccorso immediato, prestato da personale medico oppure da profani opportunamente addestrati al servizio d'urgenza, può voler dire salvezza

di Mickey S. Eisenberg, Lawrence Bergner, Alfred P. Hallstrom e Richard O. Cummins

Non c'è urgenza medica più drammatica per il singolo individuo della morte improvvisa cardiaca. Il cuore può smettere di funzionare di colpo. La vittima, che è spesso apparentemente in buona salute, ha un collasso senza alcun segno premonitore oppure, nel migliore dei casi, presenta soltanto deboli sintomi: dolore al torace, dispnea, sudorazione, nausea e senso di prostrazione. Poiché quando il cuore cessa di funzionare non giunge più sangue al cervello, la vittima perde conoscenza nel giro di qualche secondo. Con il rapido avvicinarsi della morte, si susseguono convulsioni, affanno e incontinenza.

Nel passato si nutrivano poche speranze per chi veniva colpito da morte improvvisa cardiaca. Oggi, circolazione e respirazione possono essere entrambe sostenute artificialmente in modo da poter rianimare la vittima ripristinando, con un vero e proprio shock, le normali contrazioni ritmiche del cuore. Questi interventi risultano però vani se non vengono adottati entro qualche minuto dal collasso. Ogni minuto che trascorre senza un adeguato trattamento riduce drasticamente le possibilità di rianimazione.

Poiché è necessario che l'assistenza medica giunga alla vittima il più in fretta possibile, in tutti gli Stati Uniti - così come in molti altri paesi - sono stati istituiti servizi che inviano rapidamente specialisti sul luogo da cui è giunta la richiesta di soccorso. Il personale medico d'urgenza deve essere in grado di mettersi in moto in tempi sufficientemente brevi per riuscire a strappare alla morte le vittime, permettendo loro di ritornare a una vita normale e socialmente produttiva.

Senza dubbio un'adozione più ampia di questo tipo di intervento potrebbe salvare un maggior numero di vite. Valutazioni effettuate sui programmi di medicina d'urgenza hanno mostrato che l'equipaggiare autoambulanze attrezzate in modo speciale con personale particolarmente addestrato può essere un modo efficace e funzionale per far fronte al problema della morte improvvisa cardiaca. A completamento di ciò, grandi investimenti in personale e attrezzature rappresentano iniziative attuabili con una spesa non eccessiva, che potrebbero ridurre sensibilmente la mortalità delle vittime di morte improvvisa cardiaca.

Il termine «morte improvvisa cardiaca» si riferisce all'arresto inatteso della respirazione e della circolazione, causato da una malattia cardiaca che ne è alla base, come l'aterosclerosi delle coronarie. A meno che la respirazione e la circolazione non vengano ripristinate nel giro di qualche minuto dopo l'attacco, la morte biologica permanente, coincidente con un danno cerebrale irreversibile, non potrà essere evitata.

La morte improvvisa cardiaca è un quadro importante che si può far rientrare nel più vasto problema della cardiopatia coronarica. Non si sa con esattezza quanti siano coloro che soffrono di questa malattia. Studi autoptici indicano che segni precoci di essa sono presenti in quasi tutti gli americani di sesso maschile a partire dai trent'anni. Sfortunatamente, la sola presenza di fattori che notoriamente contribuiscono al rischio di cardiopatie non permette di identificare con precisione i soggetti suscettibili di morte improvvisa cardiaca. In realtà, la maggior parte dei casi di cardiopatia coronarica non viene di solito scoperta fino a quando non insorgono gravi complicanze, come l'angina pectoris (dolore causato da una mancanza di sangue ossigenato che giunge al muscolo cardiaco), l'infarto miocardico (morte del tessuto muscolare cardiaco) o, nel peggiore dei casi, la morte improvvisa cardiaca.

Queste e altre complicanze della cardiopatia si manifestano ogni anno negli Stati Uniti sotto forma di un milione e mezzo circa di attacchi cardiaci. Circa tre quarti di coloro che subiscono un attacco cardiaco sono ricoverati in ospedale, dove sono sottoposti a una terapia coronarica avanzata. Per questi pazienti si rilevano eccellenti tassi di sopravvivenza: più dell'80 per cento dei malati ricoverati sopravvive all'attacco e viene dimesso.

Ma oltre 350 000 cittadini statunitensi all'anno non sono così fortunati. L'attacco cardiaco sopraggiunge spesso senza alcun preavviso; quando i sintomi si manifestano, di solito precedono l'attacco di meno di un'ora, lasciando alla vittima poche possibilità di essere condotta in un ospedale prima del collasso. Sono queste le vittime della morte improvvisa cardiaca. La maggioranza ha già sofferto di cardiopatia coronarica, un quadro che comprende angina pectoris e infarti miocardici pregressi. In circa il 20 per cento dei casi, tuttavia, la morte improvvisa cardiaca costituisce per la vittima la prima indicazione che cuore e coronarie sono realmente malati.

Di norma, le cavità di un cuore sano si contraggono e si rilasciano ritmicamente e la loro attività è coordinata da impulsi elettrici provenienti da due strutture, o nodi; gli impulsi sono trasmessi lungo fibre nervose incluse nel muscolo cardiaco. Nonostante che, per una quantità di ragioni, il cuore presenti molte forme di aritmia, ovvero anomalie del ritmo cardiaco, solo alcune aritmie sono associate alla morte improvvisa cardiaca. Di esse la più frequente è la fibrillazione ventricolare, che si osserva in circa due casi su tre. Durante la fibrillazione ventricolare, gli impulsi che controllano la contrazione dei ventricoli, cioè delle due grandi cavità inferiori del cuore, diventano caotici. Il muscolo si contrae spasmodicamente, senza un'attività di pompaggio coordinata e, di conseguenza, il sangue non viene espulso. Ne risulta che la pressione sanguigna e la frequenza del battito scendono immediatamente a zero.

Il meccanismo che innesca la fibrillazione ventricolare non è chiaro. Tra le possibili cause si annoverano l'ischemia del miocardio (flusso di sangue ridotto al muscolo cardiaco), squilibri elettrolitici, o alterazione delle piastrine, e stress psicologici. Solo nel 20-30 per cento dei casi di fibrillazione ventricolare l'infarto miocardico acuto è un fattore precipitante.

L'assenza di reperti costanti nei pazienti di morte improvvisa cardiaca, sottoposti a rianimazione, suggerisce la possibile esistenza di numerose e diverse sindromi eziologiche. Le morti improvvise cardiache si possono ciononondimeno suddividere in tre vasti gruppi in base alla causa immediata di alterazione del ritmo: malattie associate all'infarto miocardico, malattie associate all'ischemia miocardica e malattie associate a disturbi primari del ritmo cardiaco, di difficile interpretazione.

Indipendentemente dalle cause che determinano, o contribuiscono a determinare, la morte improvvisa cardiaca, le moderne tecniche di rianimazione d'urgenza riescono ad arrestare la morte biologica per un tempo sufficientemente lungo da poter riportare in vita il paziente. La rapidità con cui queste tecniche vengono applicate è il fattore che determina le possibilità di sopravvivenza delle vittime.

Il trattamento immediato della morte improvvisa cardiaca è la rianimazione cardiopolmonare, una serie ripetuta di atti respiratori bocca a bocca e di compressioni toraciche, che sospinge una piccola quantità di sangue ossigenato (dal 5 al 30 per cento circa della gittata cardiaca normale) verso il cervello, il cuore e altri organi corporei d'importanza vitale. La rianimazione cardiopolmonare è poi seguita da un trattamento medico definitivo, che cerca di ripristinare la circolazione e la respirazione normali. Il trattamento comprende l'inserimento nella trachea di una cannula per la respirazione, nella somministrazione di farmaci e nella defibrillazione.

La defibrillazione interrompe le contrazioni caotiche e parossistiche del muscolo cardiaco e restituisce al normale segnapas-

La prevenzione della morte improvvisa

Si intende per morte improvvisa la morte non preceduta da apprezzabili segni premonitori nel corso delle 24 ore precedenti e non dovuta a cause traumatiche. Si intende per morte improvvisa cardiaca la morte improvvisa dovuta a cause cardiache. La morte improvvisa e la morte improvvisa cardiaca sono eventi possibili in tutte le età della vita, ma i dati epidemiologici sembrano dimostrare una diversa frequenza della morte improvvisa cardiaca nei tre gruppi di popolazione tra 1 e 20 anni, tra 21 e 75-80 anni, oltre i 75-80 anni. Nel gruppo di età comprese tra 1 e 20 anni la morte improvvisa sembra dovuta prevalentemente alle infezioni, specie delle basse vie respiratorie, con shock settico; relativamente elevata appare altresì la responsabilità di affezioni quali l'epilessia e le emorragie endocraniche, mentre molto numerose (circa il 30 per cento) sono comunque le morti improvvise di cui non si riesce a trovare una convincente spiegazione. Una ricerca recente, condotta dal Department of Epidemiology dell'Università di Pittsburgh, attribuisce a cause cardiache (miocarditi, cardiomiopatie, disritmie, anomalie congenite delle coronarie, difetti congeniti delle valvole e dei setti, aneurisma aortico, eccetera) non più del 25 per cento delle morti improvvise in persone di età compresa tra 1 e 21 anni. Invece, la morte improvvisa cardiaca è indicata come la causa di gran lunga più frequente in tutte le statistiche di morte improvvisa relative alla popolazione tra i 21 e i 75-80 anni. Gli studi epidemiologici danno percentuali di morte improvvisa cardiaca tra il 75 e l'80 per cento sul totale delle morti improvvise in questi gruppi di età; e se, sempre in questi gruppi di età, insieme con i casi di morte improvvisa dovuti direttamente a malattia cardiaca, si considerano i casi di morte improvvisa dovuti presumibilmente ai farmaci usati per curare le malattie cardiache, la percentuale sale ulteriormente. La malattia cardiaca che più spesso è (direttamente o indirettamente) all'origine della morte improvvisa cardiaca nelle persone tra i 21 e 75-80 anni di età è la cardiopatia ischemica (angina pectoris e infarto miocardico); si ritiene che essa sia responsabile dell'80 per cento, e forse più, dei casi di morte improvvisa cardiaca. La causa più frequente della morte improvvisa cardiaca nelle persone affette da cardiopatia ischemica è la fibrillazione ventricolare, spesso preceduta (come suggeriscono anche alcune ricerche recentissime) da tachicardia ventricolare. La percentuale delle morti improvvise cardiache sul totale delle morti improvvise tende a diminuire sensibilmente nella vecchiaia estrema, oltre i 75-80 anni.

Così stando le cose, si deve concludere che nella grande maggioranza dei casi, morte improvvisa non è sempre sinonimo di morte imprevedibile. Al contrario. È chiaro che dobbiamo includere la morte improvvisa nel novero degli eventi probabili in un grandissimo numero di persone, e dunque degli eventi che dobbiamo «prevedere» e, se possibile, «cercare di prevenire».

La prevenzione della causa lontana della maggioranza delle morti improvvise cardiache (la cardiopatia ischemica) ha fatto grandi passi nel corso degli ultimi 20 anni e anche in Italia, come negli altri paesi progrediti - sia attraverso la correzione di errati usi alimentari e di pericolose abitudini voluttuarie, sia, soprattutto, attraverso l'efficace e sistematica terapia dell'ipertensione arteriosa e del diabete mellito - notevoli risultati sono stati ottenuti nella lotta contro l'aterosclerosi coronarica, con marcata riduzione della sua incidenza. Inoltre, alcuni difetti cardiaci che possono provocare la morte improvvisa sono ora correggibili. Ma oggi - come dimostrano i dati riportati nell'articolo di M. S. Eisenberg e collaboratori - è altresì possibile ottenere una sensibile riduzione della mortalità, intervenendo con un'apposita organizzazione di soccorso specializzato sulla più frequente causa prossima della morte improvvisa cardiaca: la fibrillazione ventricolare.

Di una tale organizzazione non esiste in Italia, su scala nazionale, nemmeno l'embrione. Si deve sperare che il Parlamento approvi rapidamente le proposte che, in materia, la Direzione generale della programmazione del Ministero della sanità intende avanzare in occasione dell'imminente presentazione del Piano sanitario nazionale. Intanto, tra le sparse iniziative locali, vale la pena di segnalarne due, che hanno preso corso nelle ultime settimane e che appaiono di notevole interesse. Una è il servizio di soccorso con elicotteri (dotati di organizzazione, di strumenti e di personale idonei a rispondere in modo efficace entro due o tre minuti dalla chiamata); nei prossimi mesi esso sarà impiegato (a cura dell'ACI e di amministrazioni locali) in Piemonte, nel Lazio e in Emilia-Romagna. L'altra è il servizio di pronto intervento specializzato, che utilizza motociclette (particolarmente adatte a districarsi velocemente nel caotico traffico delle grandi città) e in cui è impegnato, a Roma, utilizzando 18 medici specialisti o specializzandi, il Centro di rianimazione del Policlinico Umberto I.

Lucio Rosaia

si dell'attività elettrica del cuore la capacità di comandare la scansione temporale delle contrazioni. Essa viene ottenuta applicando, attraverso il torace della vittima, uno shock elettrico che depolarizzi simultaneamente tutte le cellule cardiache e quindi «azzeri» i vari segnapassi del cuore.

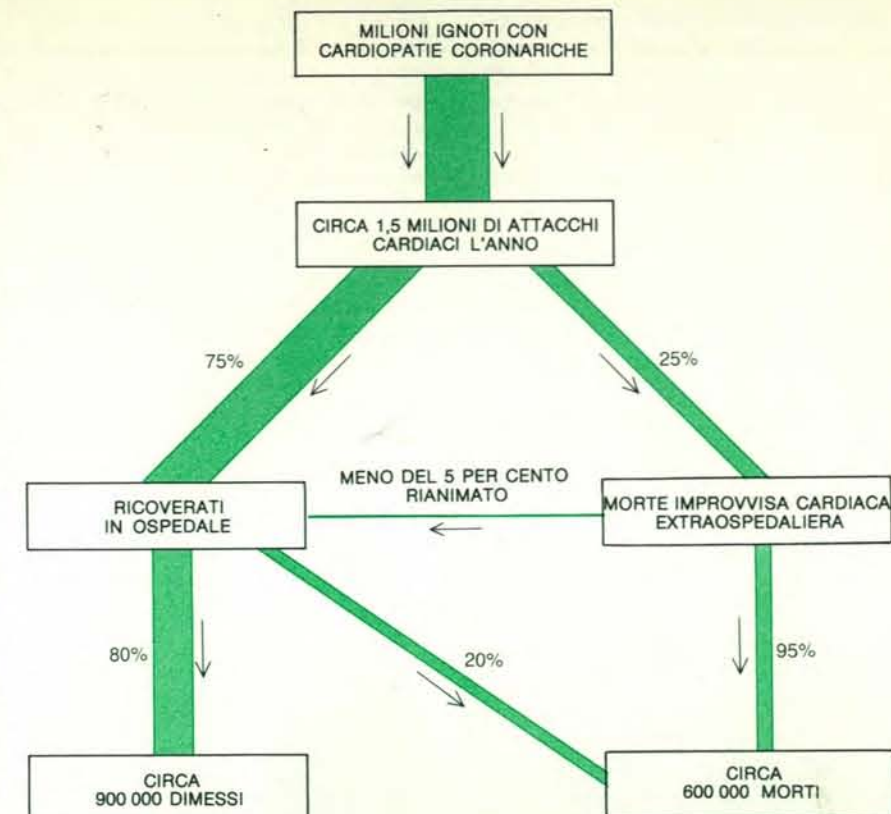
Agli inizi degli anni sessanta, i servizi pubblici di medicina d'urgenza degli Stati Uniti per questo genere di assistenza erano gravemente carenti. Infatti l'assistenza preospedaliera era fornita soprattutto da ambulanze private, scarsamente attrezzate, e da agenzie di pompe funebri, i cui veicoli funzionavano anche spesso da carri funebri e il cui personale era addestrato al pronto soccorso in modo soltanto approssimativo. I rapporti della National Academy of Sciences e della President's Commission on Highway Safety hanno confermato i livelli inadeguati dell'assistenza medica d'urgenza e la disomogeneità di competenze del personale.

La rianimazione cardiopolmonare è il trattamento immediato da praticare a una vittima di arresto cardiaco. Una volta accertato che questa ha sicuramente perduto conoscenza, le si orientano testa e collo in modo da assicurare che le vie respiratorie non siano ostruite dalla parte posteriore della lingua (*1*). Ciò si ottiene reclinando all'indietro la testa e sollevando il mento. L'assenza di respirazione è confermata auscultando, o cercando comunque di percepire, l'espirazione (*a*), e osservando se il torace della vittima si alza o si abbassa (*b*). Se il soggetto non respira, viene avviata immediatamente la respirazione bocca a bocca (*2*). Le narici vengono tenute chiuse e il soccorritore applica la propria bocca a quella della vittima, soffiando aria nei polmoni. Dopo due insufflazioni iniziali, praticate in un tempo compreso tra i quattro e i sette secondi, viene controllato il battito cardiaco rilevando le pulsazioni delle carotidi nel collo, da ciascun lato del pomo di Adamo (*3*). Se non si rileva alcuna pulsazione, si iniziano subito le compressioni toraciche (*4*). Il fondo del palmo di una mano viene appoggiato sullo sterno della vittima, a sei centimetri circa dall'incavo sternale. Si intrecciano le dita delle mani e si applica una pressione secondo un movimento a dondolo, premendo con decisione verso il basso e tenendo i gomiti in estensione. Il torace viene abbassato di circa cinque centimetri da 80 a 100 volte al minuto. Se il soccorritore opera da solo, deve alternare le compressioni toraciche alla ventilazione: ogni 15 compressioni vengono praticate due insufflazioni. Le vittime di arresto cardiaco, che vengono sottoposte a rianimazione cardiopolmonare immediatamente dopo il collasso, hanno una probabilità di sopravvivenza all'incirca doppia di quelle a cui la rianimazione non viene affatto praticata. Il procedimento è qui illustrato in modo schematico; istruzioni particolareggiate dovrebbero essere ottenute dagli organismi che sono deputati alla sanità pubblica o dai vigili del fuoco.

È stato più il desiderio di ridurre la mortalità da incidenti stradali che non quello di salvare le vittime di attacchi cardiaci a stimolare, verso la metà degli anni sessanta, miglioramenti nell'assistenza medica d'urgenza. Da un più rapido trattamento d'urgenza trassero comunque vantaggio le vittime di arresto cardiaco. Servizi medici d'urgenza sono stati incoraggiati dal Governo federale attraverso il National Highway Safety and Traffic Act del 1966. Un corso ufficiale di addestramento nell'assistenza medica d'urgenza per tecnici, della durata di 81 ore, è stato messo a punto in breve tempo per istruire sulle tecniche di base a sostegno della vita, tra cui la rianimazione cardiopolmonare.

La terapia coronarica avanzata non fu inclusa come elemento specifico nel servizio medico di emergenza fino alla seconda metà degli anni sessanta, sebbene fosse noto che più della metà dei decessi dovuti a cardiopatie avveniva fuori dall'ospedale, entro una o due ore dall'insorgenza dei sintomi. Unità mobili di terapia intensiva, in grado di far fronte a episodi di morte improvvisa cardiaca, sono state introdotte per la prima volta nel 1966 a Belfast, nell'Irlanda del Nord, da J. Frank Pantridge e J. S. Geddes del Royal Victoria Hospital. Queste unità erano dotate di un medico e di un infermiere che avessero familiarità sia con la defibrillazione e con la somministrazione di farmaci contro le aritmie, sia con il trattamento dell'infarto miocardico. In breve, anche negli Stati Uniti si sono viste circolare ambulanze simili per l'assistenza d'urgenza, ma esse, invece di essere equipaggiate con personale ospedaliero, venivano dotate di personale paramedico: profani che avevano seguito un corso di addestramento, di durata compresa tra 400 e 1500 ore, su come trattare gli arresti cardiaci al di fuori dell'ospedale, altre malattie acute e urgenze mediche. I primi programmi paramedici di questo tipo - a Seattle, Miami, Los Angeles e Columbus - sono stati istituiti rispettivamente da Leonard A. Cobb della School of Medicine dell'Università di Washington, Eugene L. Nagel della School of Medicine dell'Università di Miami, da J. Michael Criley della School of Medicine dell'Università della California a Los Angeles e da James V. Warren della School of Medicine della Ohio State University.

L'elaborazione di programmi paramedici e l'addestramento - che è oggi in continua espansione - di tecnici per l'assistenza medica d'urgenza hanno ricevuto un sostanziale appoggio nel 1973, quando il Congresso degli Stati Uniti ha varato l'Emergency Medical Service Systems Act, poi rinnovato nel 1976 e ancora nel 1979. La legislazione originaria ha stanziato 185 milioni di dollari per la messa a punto di programmi paramedici e per l'addestramento di tecnici per l'assistenza medica d'urgenza in 300 zone dell'intero paese; inoltre ha incoraggiato un modo nuovo di affrontare il problema dell'assistenza medica d'urgenza, basato su sistemi integrati.

MILIONI IGNOTI CON CARDIOPATIE CORONARICHE
CIRCA 1,5 MILIONI DI ATTACCHI CARDIACI L'ANNO
75%
25%
RICOVERATI IN OSPEDALE
MENO DEL 5 PER CENTO RIANIMATO
MORTE IMPROVVISA CARDIACA EXTRAOSPEDALIERA
80%
20%
95%
CIRCA 900 000 DIMESSI
CIRCA 600 000 MORTI

La cardiopatia coronarica si manifesta ogni anno negli Stati Uniti con circa un milione e mezzo di attacchi cardiaci, causati da ostruzione parziale o totale delle coronarie e da aritmie (disturbi del ritmo cardiaco). Circa un quarto dei casi è classificabile come morte improvvisa cardiaca al di fuori dell'ospedale: si tratta di situazioni in cui la cardiopatia coronarica ha fatto precipitare l'arresto della circolazione e della respirazione, con scarsi o nulli segni premonitori. Attualmente meno del 5 per cento delle vittime di morte improvvisa cardiaca viene rianimato con successo.

In quanto sistema, i servizi di medicina d'urgenza sono una complessa integrazione di elementi diversi: numeri telefonici per chiamate d'urgenza, personale addestrato, autoambulanze equipaggiate in modo specifico, medici addetti alla supervisione e collegamenti in telemetria, mediante i quali i dati fisiologici vitali possono essere trasmessi agli ospedali. Sebbene i servizi di medicina d'urgenza siano progettati per far fronte a tutti i tipi di emergenza medica - e non soltanto all'arresto cardiaco - la capacità di questi servizi di rispondere tempestivamente li rende una componente d'importanza decisiva di qualunque programma inteso a prestare soccorso alle vittime di morte improvvisa cardiaca.

Non più del cinque per cento delle vittime di morte improvvisa cardiaca viene oggi rianimato e dimesso dall'ospedale. Questo tasso di sopravvivenza così esiguo riflette il fatto che, nella stragrande maggioranza dei casi, il servizio medico d'urgenza arriva troppo tardi, seppure arriva. La percentuale dei sopravvissuti può essere aumentata in modo sostanziale attraverso uno sforzo concertato su scala nazionale. Poiché il numero assoluto di casi è molto grande, anche miglioramenti secondari nella tempestività con cui si iniziano la rianimazione cardiopolmonare e la defibrillazione nelle vittime di arresto cardiaco potrebbero tradursi in un grande numero di vite salvate.

Numerose circostanze, come l'età delle vittime, la presenza di testimoni al momento del collasso e la fortuna di poter essere sottoposto da un astante alla rianimazione cardiopolmonare, hanno un notevole peso sulle probabilità di ripresa da morte improvvisa cardiaca. Tali circostanze sono però imprevedibili e sono modificabili con difficoltà. Altri fattori determinanti per la sopravvivenza della vittima sono direttamente correlati con la struttura del programma di terapia cardiaca d'urgenza, e questa struttura può essere migliorata in modo significativo. Sono soprattutto le differenze nell'assistenza medica d'urgenza a spiegare l'ampia discrepanza tra le diverse comunità per quanto riguarda i tassi di sopravvivenza riportati, relativi alle morti improvvise cardiache registrate al di fuori dell'ospedale: da un minimo del 3 per cento a un massimo del 30 per cento. I fattori controllabili, che possono produrre la variazione più immediata nella percentuale di rianimazioni che hanno avuto esito positivo, sono gli intervalli di tempo tra il collasso e l'inizio della rianimazione cardiorespiratoria e tra il momento del collasso e l'avvio della terapia

definitiva (vale a dire della defibrillazione).

Le comunità hanno strutture demografiche, geografie e risorse finanziarie diverse: pertanto hanno sviluppato forme diverse di servizi di medicina d'urgenza. Questi servizi possono essere forniti o da unità normali con tecnici per l'assistenza medica d'urgenza, o da unità specializzate, attrezzate con defibrillatori per intervenire *in loco* ed equipaggiate con personale paramedico, o dalla combinazione delle due.

Non sorprende che gli studi condotti abbiano messo in luce il fatto che nelle comunità dotate soltanto di unità di base si riscontrino, nei casi di attacchi cardiaci improvvisi, bassi tassi di sopravvivenza. Poiché i tecnici per l'assistenza medica d'urgenza possono attuare sul luogo dell'evento soltanto la rianimazione cardiopolmonare e devono poi trasportare la vittima a un ospedale per una terapia più avanzata, ci vogliono più di 20 minuti prima che la vittima di morte improvvisa cardiaca sia in grado di essere sottoposta a defibrillazione.

Pertanto, i programmi che forniscono i tecnici di defibrillatori migliorano in misura significativa il tasso di sopravvivenza delle vittime di morte improvvisa cardiaca. In un sobborgo di Seattle, che non disponeva di altri sistemi avanzati di assistenza d'urgenza, l'aggiunta di defibrillatori alle unità equipaggiate con tecnici per l'assistenza medica d'urgenza ha determinato un incremento di quasi cinque volte dei tassi di sopravvivenza. Inoltre, uno studio effettuato sulle comunità rurali dello Iowa ha mostrato che questa sola iniziativa si è tradotta in analoghi vantaggi.

I programmi di medicina d'urgenza che si basano esclusivamente su personale paramedico realizzano risultati migliori di quelli dei programmi che si fondano esclusivamente sull'impiego di tecnici. Tuttavia, il costo per un addestramento più completo e per un equipaggiamento più specializzato (comprendente l'impiego di defibrillatori) ne ha limitato il numero.

I tassi di sopravvivenza più elevati si registrano in quelle comunità che impiegano, per l'assistenza medica d'urgenza, un sistema combinato di tecnici e di personale paramedico, come accade a Seattle, a Miami e a Columbus. I tecnici arrivano tre o quattro minuti dopo il collasso e avviano la rianimazione cardiopolmonare; diversi minuti più tardi sopraggiunge il personale paramedico che fornisce un'assistenza più avanzata, in particolare pratica la defibrillazione. Tassi di sopravvivenza ancora più elevati si raggiungono se anche i tecnici per l'assistenza medica d'urgenza vengono dotati di defibrillatori. Le unità di cui essi fanno parte possono allora iniziare la rianimazione cardiopolmonare e praticare la defibrillazione qualora le condizioni della vittima giustifichino un trattamento di questo genere. Quando, trascorso un certo tempo, sopraggiunge il personale paramedico, esso può somministrare farmaci, inserire cannule nelle vie respiratorie e praticare tanti shock defibrillatori quanti sono necessari. Un programma combinato di questo tipo, avviato nel 1979 presso la Washington's King County, è stato coronato da successo: il tasso di sopravvivenza, nei casi extraospedalieri di fibrillazione ventricolare in cui sono trascorsi più di quattro minuti tra l'arrivo dei tecnici e quello del personale paramedico, è aumentato dal 18 al 38 per cento.

TEMPO DI RISPOSTA DA PARTE DEL PERSONALE PARAMEDICO (MINUTI)	ALTRA ARITMIA				FIBRILLAZIONE VENTRICOLARE			
	COLLASSO SENZA TESTIMONI		COLLASSO CON TESTIMONI		COLLASSO SENZA TESTIMONI		COLLASSO CON TESTIMONI	
	NESSUNA RIANIMAZIONE DA PARTE DEGLI ASTANTI	RIANIMAZIONE DA PARTE DEGLI ASTANTI	NESSUNA RIANIMAZIONE DA PARTE DEGLI ASTANTI	RIANIMAZIONE DA PARTE DEGLI ASTANTI	NESSUNA RIANIMAZIONE DA PARTE DEGLI ASTANTI	RIANIMAZIONE DA PARTE DEGLI ASTANTI	NESSUNA RIANIMAZIONE DA PARTE DEGLI ASTANTI	RIANIMAZIONE DA PARTE DEGLI ASTANTI
6-10	0	0	0	0,01	0	0,01	0,1	0,3
2-6	0	0	0,01	0,1	0	0,1	0,3	0,5
0-2	0	0,01	1	0,3	0,1	0,3	0,5	0,7

La possibilità di successo nella rianimazione (*rettangoli in colore*) praticata a una vittima di morte improvvisa cardiaca dipende in modo critico dalle particolari circostanze in cui ha avuto luogo l'arresto cardiaco e dalla tempestività con cui l'assistenza definitiva è stata erogata da personale medico o paramedico. La natura dell'aritmia, che si traduce per la vittima nell'arresto della circolazione sanguigna, e la presenza di persone che assistono al conseguente collasso non possono essere controllate. È possibile migliorare il tempo di risposta delle unità paramediche e la probabilità che gli astanti siano in grado di praticare la rianimazione cardiopolmonare, se si aumenta sia il numero di autoambulanze a disposizione sia il numero di persone destinate all'assistenza medica d'urgenza, insegnando anche ai profani la tecnica di rianimazione cardiopolmonare.

	INFARTO MIOCARDICO	ISCHEMIA MIOCARDICA	ALTERAZIONE PRIMARIA DEL RITMO
PERCENTUALE DI MORTI IMPROVVISE CARDIACHE	20-30	?	?
PROVE RACCOLTE IN CORSO DI AUTOPSIA	OCCLUSIONE DELLE CORONARIE; GRAVE CORONAROPATIA	CORONAROPATIA	CUORE IN APPARENZA NORMALE IN POCHI CASI, ALTRI CASI DI CORONAROPATIA
DURATA DEI SINTOMI PREMONITORI PRECEDENTI IL COLLASSO	DA MINUTI A ORE	MINUTI	SECONDI; NESSUN SEGNO PREMONITORE
PROGNOSI A LUNGO TERMINE DOPO UNA RIANIMAZIONE RIUSCITA	FAVOREVOLE	FAVOREVOLE	SFAVOREVOLE

Tre cause di disfunzioni cardiache che portano alla morte improvvisa cardiaca sono l'infarto miocardico (necrosi del muscolo cardiaco per mancanza di ossigeno), l'ischemia (ridotto rifornimento di ossigeno e di sostanze nutritive al cuore) e disturbi a carico del ritmo cardiaco, di difficile interpretazione. Le morti improvvise cardiache sono raggruppate sulla base di queste cause.

In generale, l'aggiunta di altre autoambulanze, di un equipaggiamento sussidiario e di personale addestrato abbrevierà il tempo che intercorre tra il collasso e l'inizio della rianimazione cardiopolmonare come pure il tempo che intercorre tra il collasso e la somministrazione della terapia definitiva. Da un certo punto in poi, però, si assiste a una diminuzione dei vantaggi collaterali: ulteriori miglioramenti nel tempo di risposta richiedono sostanziali investimenti in autoambulanze e in personale, con scarso o nullo miglioramento dei tassi di sopravvivenza.

Un tempo di risposta medio di quattro minuti è stato realizzato, per esempio, utilizzando per l'assistenza d'urgenza 100 veicoli dei vigili del fuoco, distribuiti sui circa 1300 chilometri quadrati della King County. Abbiamo calcolato che l'abbassamento di questo tempo di altri 30 secondi richiederebbe un'ulteriore aggiunta di 31 veicoli, con un costo annuo per veicolo di circa 43 000 dollari (facendo l'ipotesi che non sia necessario altro personale paramedico). L'aggiunta di personale paramedico è ancora più costosa: nella King County il mantenimento di un'unità paramedica operativa costa all'anno 500 000 dollari. In una situazione del genere un'alternativa vantaggiosa dal punto di vista economico sarebbe migliorare la qualità delle unità di base che sono equipaggiate con tecnici per l'assistenza medica d'urgenza dotandole di defibrillatori. Poiché non sarebbe richiesto altro personale, una più tempestiva applicazione della defibrillazione potrebbe essere ottenuta a un costo relativamente modesto.

Il costo non è il solo fattore che può rendere proibitivo l'aumento del numero di unità di assistenza medica d'urgenza, dotate di tecnici o di personale paramedico. In molte comunità, la geografia della regione in cui esse vivono, l'altezza degli edifici e la rete stradale pongono limiti ai tempi di risposta, che non possono sostanzialmente essere abbassati, indipendentemente dal numero delle unità aggiunte. Esistono, comunque, altre strategie che possono abbreviare il tempo di risposta all'assistenza d'urgenza: alcune sono rapidamente attuabili e poco costose.

Per esempio, rendendo semplicemente più facile l'accesso della gente ai centri dove può segnalare una situazione d'emergenza e aumentando l'efficienza degli smistamenti si possono risparmiare secondi preziosi nel tempo di risposta. Molte comunità hanno perciò istituito un numero telefonico facile da ricordare (negli Stati Uniti esso è il 911), attraverso il quale è possibile riferire tutte le situazioni d'emergenza. Inoltre, sono stati sviluppati sistemi di smistamento, assistiti da calcolatore: essi possono rapidamente identificare gli indirizzi con estrema precisione e individuare le più vicine autoambulanze disponibili per l'assistenza d'urgenza.

Un'altra misura che offre qualche vantaggio, con investimenti relativamente ridotti, è l'istituzione di programmi di addestramento per la rianimazione cardiopolmonare. Un programma del genere fu messo a punto per la prima volta nel 1971 da Cobb a Seattle. Da allora numerose altre comunità in tutto il paese hanno istituito programmi analoghi, con un addestramento fornito sia come servizio pubblico, sia dietro pagamento di una piccola quota. Si suppone così che un grande numero di persone addestrate, sparse in tutta la comunità, possa aumentare le possibilità che una vittima sia tempestivamente sottoposta a rianimazione cardiopolmonare. Le vittime che, subito dopo il collasso, vengono tempestivamente sottoposte a questa pratica hanno una possibilità di sopravvivenza due volte superiore alle altre. La American Heart Association e la Croce Rossa americana hanno dato un ampio sostegno a questi tentativi di addestramento. I corsi, che durano dalle tre alle nove ore, sono anche un'opportunità ideale per far conoscere ai partecipanti i numeri telefonici di emergenza, i fattori di rischio della cardiopatia coronarica, l'assistenza sanitaria preventiva e i segni premonitori dell'infarto miocardico.

Tuttavia, le istruzioni per la rianimazione cardiopolmonare, impartite a tutta la comunità, hanno ottenuto soltanto un successo limitato. La nostra esperienza nell'area della King County è chiarificatrice: sebbene la percentuale della popolazione addestrata sia aumentata sei volte nello scorso decennio, la percentuale di casi di arresto cardiaco, in cui la rianimazione è stata impartita dagli astanti, si è stabilizzata negli ultimi anni attorno a un 30 per cento abbastanza costante. Un motivo risiede nel fatto che coloro che vogliono essere addestrati per la rianimazione cardiopolmonare sono in maggioranza giovani (età media 33 anni) e sono ugualmente suddivisi tra maschi e femmine. Per la maggior parte, le vittime di arresto cardiaco sono, invece, uomini di oltre cinquant'anni, che di solito subiscono l'attacco cardiaco a casa. Le loro mogli, in genere anch'esse oltre la cinquantina, sarebbero dunque le persone più adatte per mettere in pratica le istruzioni nel campo della rianimazione cardiopolmonare.

Evidentemente occorre studiare nuovi modi di raggiungere coloro che hanno maggiori probabilità di venirsi a trovare in

presenza di una vittima di arresto cardiaco. Recentemente sono state introdotte alcune alternative ai corsi tradizionali di rianimazione cardiopolmonare, nel tentativo di dare al maggior numero di persone possibile alcune informazioni su cosa fare in caso di morte improvvisa cardiaca. Le dimostrazioni sono state date durante incontri di baseball trasmessi per televisione su tutto il territorio nazionale; sono stati anche progettati per la televisione, come servizio pubblico, spot in più puntate. È stato proposto da tempo che le famiglie dei pazienti affetti da cardiopatia coronarica siano classificate come principali destinatari dell'addestramento alla rianimazione cardiopolmonare. Nonostante questi lodevoli sforzi e le campagne nazionali della Heart Association e della Croce Rossa americana, solo una piccola percentuale di coloro che hanno più probabilità di trovarsi in presenza di una vittima di arresto cardiaco ha ricevuto una preparazione adeguata.

Un'alternativa, che indirizza le informazioni sulla rianimazione cardiopolmonare a coloro che potrebbero applicarla nel modo più tempestivo, è stata adottata nel 1974 dai vigili del fuoco di Phoenix, in Arizona: si è trattato di istruzioni telefoniche sulla rianimazione da parte degli organizzatori del servizio medico d'urgenza. Chi segnala un caso di arresto cardiaco si trova nella condizione migliore per portare soccorso immediato alla vittima. Se chi riceve la segnalazione riesce a guidare chi chiama passo per passo attraverso l'intero procedimento della rianimazione, mentre sta per giungere l'assistenza medica professionale, può essere risparmiato tempo prezioso. Sul programma di Phoenix non si sono avuti apprezzamenti ufficiali; tuttavia, vari resoconti hanno confermato che, con il telefono, è stato possibile fornire una guida adeguata.

William B. Carter e collaboratori della King County Emergency Medical Services Division, presso il Seattle-King County Health Department, hanno messo a punto un messaggio di istruzioni per la rianimazione cardiopolmonare che può essere trasmesso per telefono a chi ha chiesto soccorso per un caso di arresto cardiaco. Tutti coloro che, alla King County, ricevono una richiesta di soccorso sono addestrati a fornire istruzioni. Il messaggio che trasmettono è in due parti: all'interlocutore viene prima detto come eseguire la respirazione artificiale; quindi, quando egli ritorna al telefono, gli vengono impartite le istruzioni sulla compressione toracica. Il messaggio è dettato con un linguaggio semplice e diretto per dare un quadro preciso dei movimenti da compiere.

Negli oltre 400 casi, che sono stati studiati fino a oggi, di rianimazione cardiopolmonare effettuata attraverso istruzioni trasmesse per via telefonica non sono stati osservati gravi effetti collaterali, né sotto forma di danni connessi con la rianimazione (e valutati sulla base di un esame dei reperti medici e autoptici), né sotto forma di impatto psicologico sull'interlocutore reclutato (valutato mediante interviste approfondite). Pare che, indipendentemente dall'esito, l'assistenza in favore di una vittima di arresto cardiaco inneschi in questa un effetto psicologico positivo. I programmi telefonici di addestramento alla rianimazione cardiopolmonare sono stati approvati di recente dalla American Heart Association e sono stati istituiti su tutto il territorio degli Stati Uniti; altrettanto è stato fatto in Inghilterra e nei Paesi Bassi.

Non c'è dubbio che, quanto più precocemente si avvia la rianimazione, tanto maggiori sono le possibilità di sopravvivenza della vittima. Tuttavia, la rianimazione cardiopolmonare non rappresenta altro che un sussidio temporaneo; la vittima ha bisogno di quell'assistenza definitiva che il personale medico e paramedico può offrire per una maggior probabilità di sopravvivenza. La defibrillazione è soltanto una delle componenti dell'assistenza definitiva, ma è probabilmente la più importante; quando è applicata tempestivamente, può rappresentare tutto ciò che è necessario per salvare la vita della vittima. Pertanto una maggior disponibilità generale di defibrillatori nella comunità può tradursi in un deciso aumento dei tassi di sopravvivenza.

La morte improvvisa cardiaca, associata alla fibrillazione ventricolare, di solito non ha luogo se sono disponibili ambulanze con personale paramedico o se il paziente si trova in ospedale, dove i defibrillatori costituiscono una dotazione standard. Essa avviene soprattutto in casa. Non si può purtroppo dotare di defibrillatori convenzionali le abitazioni dei pazienti affetti da cardiopatia coronarica perché l'uso di questi strumenti richiede la capacità di interpretare i ritmi cardiaci, una attitudine che si può acquisire solo attraverso un addestramento particolare. La necessità di un addestramento esauriente e il potenziale pericolo che corre chi opera con un defibrillatore in caso di incidente hanno confinato la defibrillazione nell'ambito specialistico.

PERSONALE DELLA AMBULANZA	AMBULANZE CHE RISPONDONO ALLA CHIAMATA D'URGENZA	TEMPO MEDIO TRA COLLASSO E RIANIMAZIONE (MINUTI)	TEMPO MEDIO TRA COLLASSO E DEFIBRILLAZIONE (MINUTI)	TEMPO MEDIO TRA COLLASSO E INTERVENTO CON TECNICHE AVANZATE (MINUTI)	TASSO DI RIANIMAZIONI RIUSCITE (PERCENTUALE)
TECNICI	1	4	20 IN OSPEDALE	20 IN OSPEDALE	3
TECNICI CON DEFIBRILLATORE	1	4	4	20 IN OSPEDALE	20
PERSONALE PARAMEDICO	1	8	8	8	10-20
TECNICI E PERSONALE PARAMEDICO	2	4	8	8	15-20
TECNICI CON DEFIBRILLATORE E PERSONALE PARAMEDICO	2	4	4	8	20-30

L'assistenza medica d'urgenza in caso di attacchi cardiaci può essere erogata da tecnici ordinari addestrati a questo tipo di assistenza, da tecnici specializzati, addestrati anche all'uso di defibrillatori, o da personale paramedico. Un tecnico ordinario segue un corso di addestramento di 81 ore e apprende le tecniche di base per il sostegno delle funzioni vitali (soprattutto la rianimazione cardiopolmonare); un tecnico specializzato ha a disposizione altre dieci ore di istruzione sull'uso dei defibrillatori. Il personale paramedico deve essere in grado di fornire un sostegno più avanzato alle funzioni vitali, tra cui la somministrazione di farmaci e l'intubazione endotracheale; pertanto il suo addestramento richiede dalle 400 alle 1500 ore. Un servizio medico d'urgenza consiste in varie combinazioni di questi diversi specialisti dell'assistenza sanitaria; ciascun gruppo è destinato ad ambulanze equipaggiate in modo specifico. Il costo elevato di un'unità paramedica confrontato con quello di un'unità dotata soltanto di tecnici significa che una comunità può permettersi un numero di unità paramediche inferiore a quello delle unità dotate soltanto di tecnici e il personale paramedico raggiunge di solito la vittima di un arresto cardiaco con maggiore ritardo. Poiché, quanto più tempestivamente la vittima di un arresto cardiaco è sottoposta a trattamento medico (sia di base sia avanzato), tanto maggiori sono le possibilità di rianimarla, i programmi di medicina d'urgenza che associano i tecnici, specializzati e no, con personale paramedico hanno dimostrato di essere i più efficaci.

Tuttavia, defibrillatori automatici di recente costruzione potranno presto cambiare la situazione. Questi strumenti sono semplici da usare: un soccorritore applica gli elettrodi adesivi al torace della vittima e la macchina fa praticamente tutto il resto. Essi sono programmati per scoprire e per riconoscere la fibrillazione ventricolare, per caricare i condensatori con la corrente adatta e quindi per produrre uno shock defibrillatorio. Una voce sintetizzata o messaggi che compaiono su un visualizzatore a cristalli liquidi aiutano l'operatore a preparare lo strumento. Anche i parenti delle eventuali vittime possono imparare ad applicare gli elettrodi e a usare lo strumento. Vari studi suggeriscono che il defibrillatore automatico è altamente efficiente: in più del 95 per cento dei casi la fibrillazione ventricolare è stata correttamente identificata. In nessun caso lo shock defibrillatorio è stato applicato in assenza di fibrillazione.

Il mettere a disposizione delle famiglie dei pazienti affetti da cardiopatia coronarica defibrillatori portatili, di facile uso, potrebbe salvare migliaia di vite all'anno, ma l'opzione non è realistica. Per la maggior parte delle comunità essa è improponibile a causa del costo. (Lo strumento ha un prezzo dai 4500 ai 7000 dollari, anche se con un aumento delle richieste se ne potrebbero ottenere di più a buon mercato.) Una misura del genere non offrirebbe, inoltre, alcun vantaggio alle vittime potenziali di morte improvvisa cardiaca non consapevoli della propria malattia.

Un modo più pratico di affrontare il problema potrebbe essere quello di collocare in luoghi pubblici alcuni defibrillatori automatici e di addestrare al loro uso i membri della comunità. Defibrillatori sugli aerei e sulle navi, negli uffici, negli edifici residenziali e nelle strutture pubbliche potrebbero avere un impatto significativo sulla tempestività con cui le vittime sono soccorse. Quali sono i vantaggi e i costi stimati? Chi pagherà? Ogni quanto tempo la gente dovrebbe essere riaddestrata al loro uso? Un profano può usare uno strumento del genere in una emergenza reale?

Un'ultima alternativa riduce al minimo il tempo per la defibrillazione: si tratta dell'impianto chirurgico di defibrillatori automatici. Mieczyslaw Mirowski e collaboratori della School of Medicine della Johns Hopkins University hanno messo a punto uno strumento che controlla di continuo il ritmo cardiaco, identifica la fibrillazione ventricolare e produce scariche defibrillatorie a bassa energia, il tutto in un tempo tra i 15 e i 20 secondi. Nel 1983 sono stati impiantati defibrillatori di questo tipo in oltre 220 pazienti e sono state realizzate con successo numerose inversioni della fibrillazione ventricolare. Questa alternativa è però raccomandabile soltanto a una piccola percentuale di pazienti ben identificabili e a rischio di morte imminente per fibrillazione ventricolare.

Numerose scelte in favore della medicina d'urgenza, a dispetto del fatto che possono effettivamente diminuire la mortalità da morte improvvisa cardiaca, sono solo una soluzione parziale del problema. La soluzione definitiva è la prevenzione della cardiopatia. Un mezzo efficace per combattere l'aterosclerosi, la causa principale della cardiopatia, è ancora da ricercare. Un cambiamento nello stile di vita e nella dieta, un maggiore esercizio fisico, il controllo dell'ipertensione e l'impiego di nuovi farmaci sono tutti elementi in grado di fornire un aiuto, ma per il momento gli sforzi devono essere concentrati verso una più rapida erogazione del trattamento medico alle vittime della morte improvvisa cardiaca. Se le tecniche di rianimazione che hanno già dimostrato la propria efficacia (soprattutto la rianimazione cardiopolmonare e la defibrillazione praticate con celerità) potranno essere attuate in modo più diffuso, sia mediante tecnici per l'assistenza medica d'urgenza, sia mediante personale paramedico o con altri programmi di sanità pubblica, ogni anno potranno essere salvate migliaia di vite in più.

a
NODO DEL SENO
NODO ATRIOVENTRICOLARE
b
c
SHOCK DEFIBRILLATORIO

La fibrillazione ventricolare è l'aritmia più frequentemente associata con la morte improvvisa cardiaca. Normalmente l'impulso elettrico che provoca la contrazione delle cellule muscolari cardiache (*a*) ha origine nel nodo del seno, raggiunge il nodo atrioventricolare e quindi si diffonde simmetricamente attraverso le pareti muscolari dei due ventricoli mediante fibre nervose incluse in esse. Come riflette l'uniforme distribuzione dei complessi elettrici normali, registrati in un elettrocardiogramma, le cavità di un cuore sano sono sottoposte a contrazioni coordinate e ritmiche. Durante la fibrillazione ventricolare (*b*), impulsi elettrici sporadici sono emessi da molteplici zone cardiache. L'elettrocardiogramma di un cuore in fibrillazione indica che le contrazioni non sono coordinate, il che si traduce in una perdita immediata della pressione sanguigna e del polso. La fibrillazione ventricolare si può arrestare con una scarica elettrica attraverso la parete toracica e il muscolo cardiaco (*c*). La scarica depolarizza tutte le cellule cardiomuscolari e permette il ripristino della naturale attività ritmica dei segnapassi. Nuovi dispositivi sono in grado di riconoscere la traccia elettrocardiografica della fibrillazione e di emettere automaticamente scariche per contrastarla; possono così mettere in grado il profano di applicare questa misura salvavita.

I glicosfingolipidi

Sono molecole della membrana cellulare la cui composizione si modifica quando la cellula si differenzia o diventa neoplastica: sfruttando questi cambiamenti è possibile perfezionare la diagnosi e la terapia del cancro

di Sen-itiroh Hakomori

Nel 1951 una donna di 66 anni fu ricoverata presso l'ospedale di Charlottesville, nello stato della Virginia, per un intervento di carcinoma allo stomaco. Aveva gruppo sanguigno di tipo *0*, quindi era un donatore universale, ma le prove di laboratorio dimostrarono anche che nel suo siero erano presenti anticorpi che avrebbero ucciso tutti i tipi di cellule ematiche a esclusione di quelli della paziente. Non era possibile trovare donatori il cui sangue fosse compatibile con i suoi anticorpi e, pertanto, la potenziale necessità di una trasfusione di sangue nell'imminenza dell'intervento la esponeva a un considerevole rischio. Per valutare questo rischio le venne praticata una piccola trasfusione di prova di circa 25 centimetri cubi di sangue di tipo *0*.

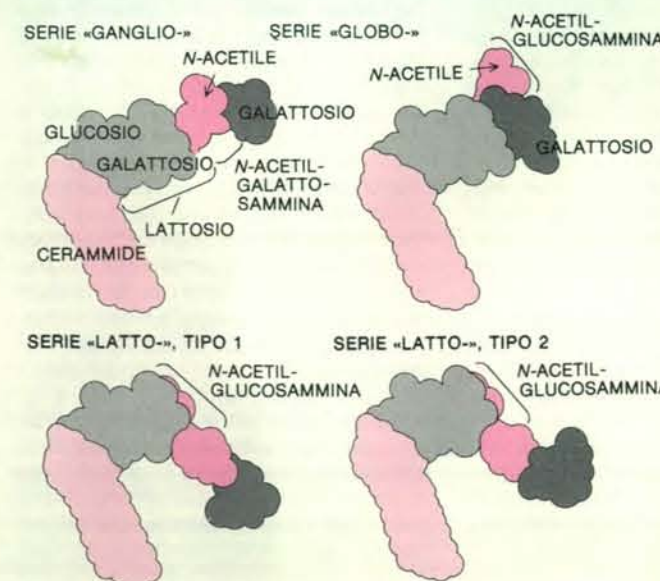
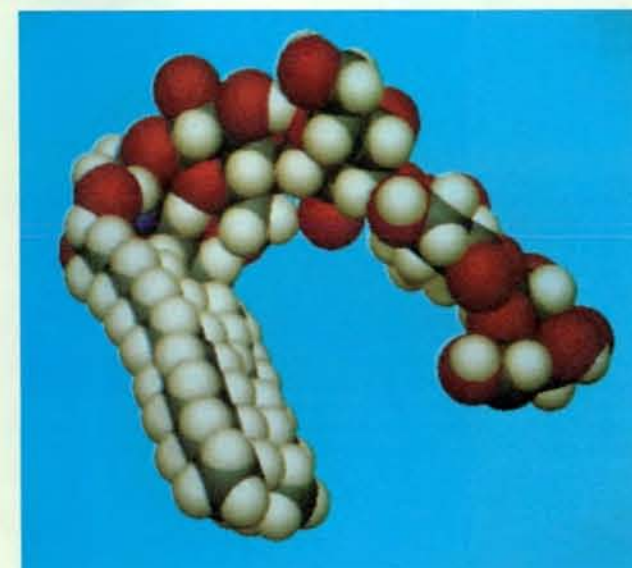
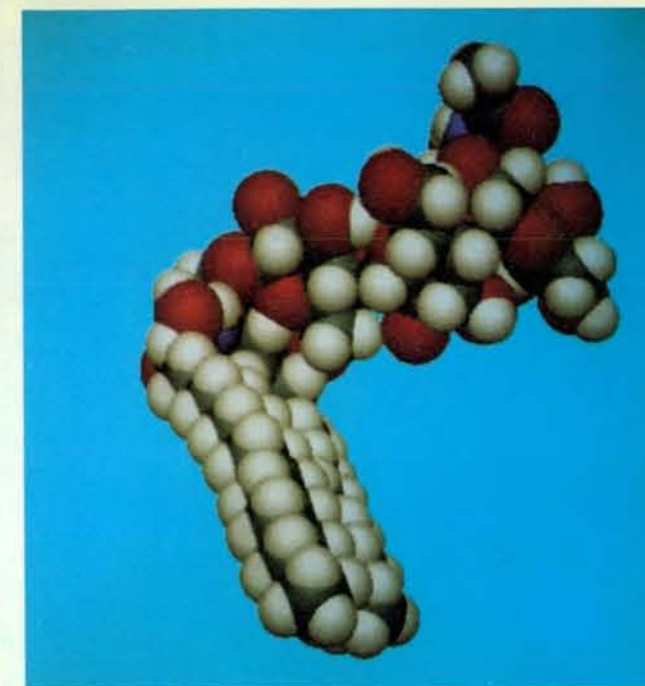
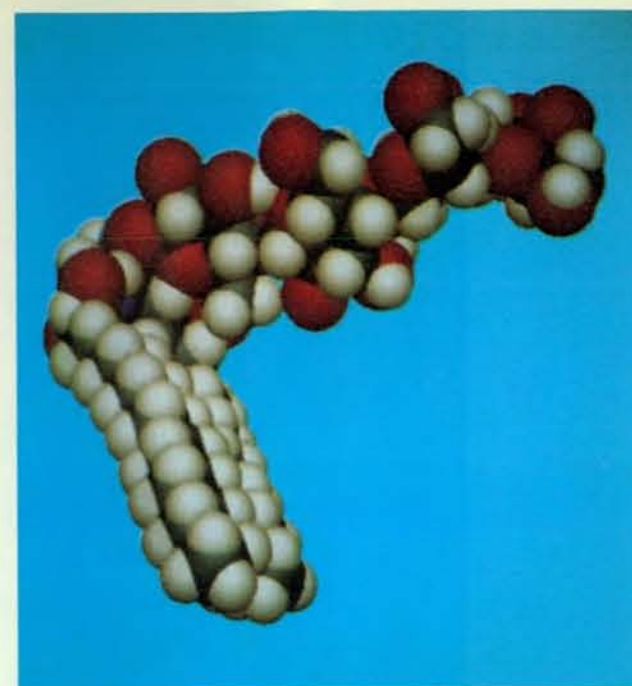
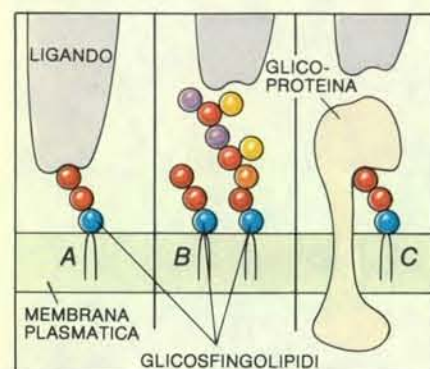
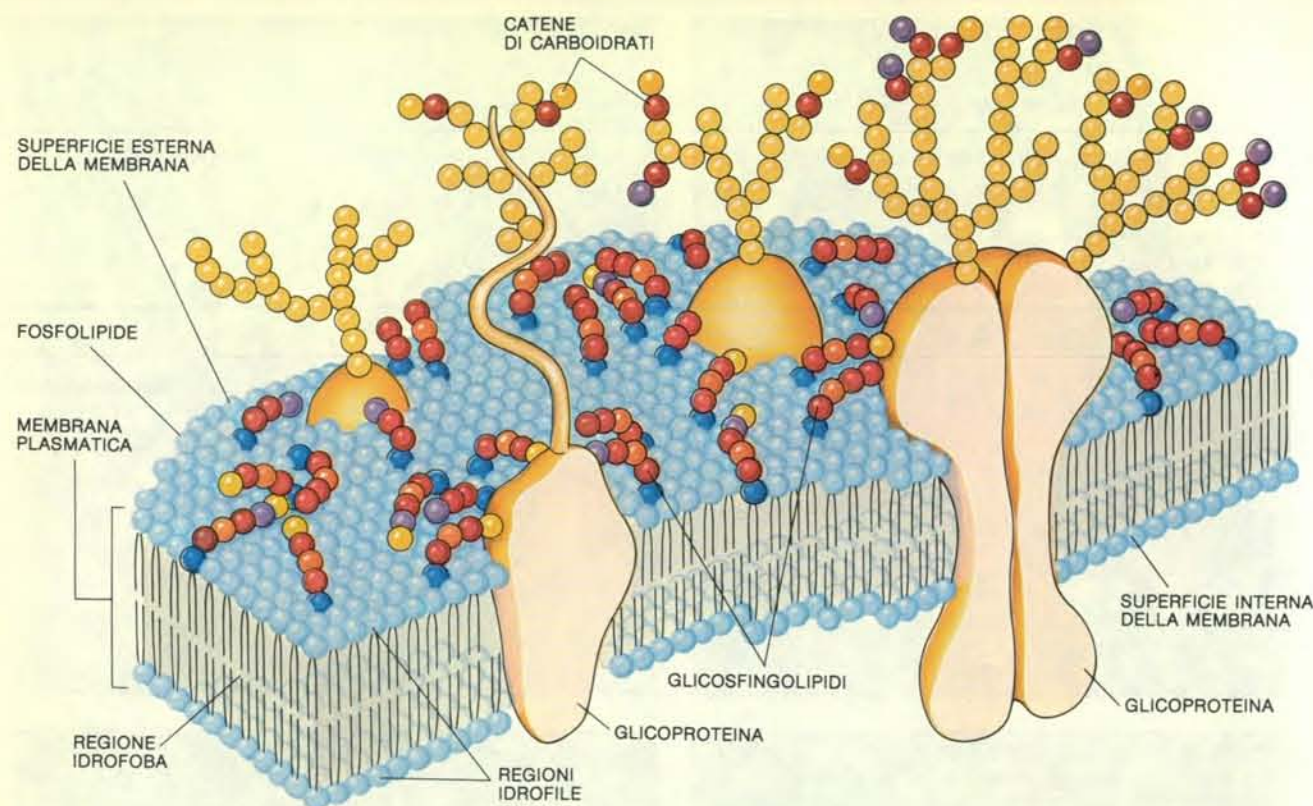
Come era prevedibile, la reazione immunitaria iniziale della donna fu drammatica. La concentrazione degli anticorpi nel siero in grado di reagire con il sangue del donatore si elevò a una parte su 512. I risultati della prova orientarono la scelta dell'intervento chirurgico: solo una parte dello stomaco poteva venir rimossa senza rischio e pertanto un poco di tessuto tumorale doveva rimanere nell'organismo. Però, con grande sorpresa e sollievo di tutti, il restante tessuto tumorale scomparve dopo l'intervento. La donna visse fino all'età di 88 anni senza più alcun segno di cancro.

Qual era stata la causa della sua sorprendente guarigione? I primi indizi li scoprì subito dopo l'operazione Philip Levine della Ortho Research Foundation. La tipizzazione di routine del sangue viene effettuata solo per il sistema di gruppi sanguigni *ABH*, nel quale sangue e tessuti sono classificati in base ai tipi clinicamente familiari *0*, *A*, *B* e *AB*. Vi sono, tuttavia, molti altri sistemi di gruppi sanguigni e l'espressione o la soppressione di vari antigeni appartenenti a ogni sistema determina un tipo sanguigno indipendente e distinto. All'interno del sistema di gruppi sanguigni *P*, la donna era la prima persona sulla quale esistesse una documentazione e i cui tessuti normali non esibissero due tipi di marcatori immunologici specifici, gli antigeni P e P_1. Una simile assenza di marcatura immunologica è la caratteristica che definisce all'interno del sistema P un raro tipo sanguigno, noto come gruppo sanguigno p, che da allora è stato trovato nella popolazione generale solo in un soggetto su 100 000. Analizzato il siero della donna, Levine fece un'altra sorprendente scoperta: esso conteneva come anticorpi delle immunoglobuline G contro gli antigeni sia P sia P_1. La formazione di quegli anticorpi doveva essere stata indotta dagli antigeni P e P_1. Da dove essi provenivano? Secondo Levine, probabilmente essi venivano espressi dal tessuto tumorale.

Trent'anni dopo l'intervento, nel nostro laboratorio è stato analizzato un campione del tessuto tumorale della donna, che era stato essiccato per congelamento e conservato nel laboratorio di Levine. L'analisi ha dimostrato che quel tessuto conteneva in effetti due tipi di antigene: uno che reagiva con l'anticorpo anti-P e l'altro che reagiva con l'anticorpo anti-P_1. Abbiamo inoltre trovato che entrambi gli antigeni sono membri di una classe di molecole, i glicosfingolipidi, incluse nelle membrane cellulari in tutto il corpo. Come indica il nome, i glicosfingolipidi sono in parte lipidi, ossia sostanze solubili nei grassi, e in parte zuccheri (glicidi). L'infisso *-sfingo-* li identifica come una classe speciale di glicolipidi, i quali includono nella loro struttura il lipide sfingosina. In questo modo abbiamo confermato la conclusione di Levine, secondo la quale la formazione di anticorpi nel siero della donna era stata indotta da antigeni incompatibili presenti nel tumore.

La violenta reazione immunitaria della donna al sangue del donatore aveva avuto luogo grazie alla presenza di anticorpi anti-P e anti-P_1 nel suo sistema, anticorpi che poteva tollerare solamente perché i suoi tessuti sani non contenevano antigeni P e P_1. Quando essa, attraverso la trasfusione, ha ricevuto antigeni P e P_1, questi hanno stimolato una rapida produzione di altri anticorpi anti-P e anti-P_1, il cui elevato livello ha presumibilmente innescato una reazione complessa da parte delle cellule del sistema immunitario; questa reazione ha distrutto selettivamente le cellule tumorali che esibivano gli antigeni incompatibili P e P_1.

Notoriamente, i glicosfingolipidi sono presenti in tutte le cellule animali e in alcune cellule vegetali. Tuttavia, ancora recentemente era stata posta, da parte dei ricercatori, una scarsa attenzione a queste molecole in quanto le loro funzioni biologiche non erano ancora state chiaramente definite. Gli studi degli ultimi 10 anni hanno inconfutabilmente stabilito che gli antigeni dei gruppi sanguigni e molti altri antigeni che vengono modificati o espressi in modo non appropriato nel corso dello sviluppo di tumori maligni sono glicosfingolipidi. Di conseguenza, i glicosfingolipidi sono diventati oggetto di intensa indagine da parte degli immunologi interessati al cancro e si sono ricavate moltissime informazioni sia sulle loro funzioni nella vita della cellula, sia sul loro ruolo nel cancro e in altre malattie.

I glicosfingolipidi, grazie alla loro distribuzione sulla membrana cellulare, regolano le interazioni della cellula sana con l'ambiente in molti modi. Per esempio, servono come marcatori che permettono di distinguere le cellule di vari organi di un animale e mediano, in questo modo, il riconoscimento e la comunicazione da cellula a cellula. Inoltre, l'espressione dei glicosfingolipidi sulla superficie cellulare si modifica allorché la cellula si divide e si differenzia, il che suggerisce che essi sono essenziali per una crescita e uno sviluppo sistematici dell'organismo. È noto che essi regolano la crescita in almeno due modi: avvertendo la densità di cellule analoghe nelle loro vicinanze, possono promuovere o inibire la divisione cellulare; interagendo con le proteine recettore presenti sulla superficie cellulare, possono anche inibire la risposta della cellula ai fattori di crescita che si potrebbero trovare nel mezzo ambiente che la circonda.

I modelli molecolari mostrano le quattro strutture centrali fondamentali dei glicosfingolipidi, una classe di molecole generalmente incluse nella membrana cellulare. Tutti i glicosfingolipidi constano di una catena di carboidrati o glicidi, unita ad angolo retto con un lipide, il cerammide. La maggior parte dei glicosfingolipidi, ma non tutti, ha in comune una di queste quattro strutture, che vengono definite in base all'identità e al legame chimico dei glicidi più vicini al cerammide. In ciascuna delle quattro strutture il cerammide è legato al glucosio e il glucosio al galattosio. Nei glicosfingolipidi della serie «ganglio-» (*in alto a sinistra*) la catena continua con un glicide di nome *N*-acetilgalattosammina, a cui si attacca un altro residuo di galattosio. Il galattosio interno e il galattosio terminale sono ciascuno legato a un glicide acido, l'acido sialico (non mostrato). Nella serie «globo-» la catena iniziale, fatta di galattosio e glucosio, si prolunga con un altro residuo di galattosio e con *N*-acetilgalattosammina (*in alto a destra*). Nella serie «latto-» la catena iniziale prosegue con un glicide, la *N*-acetilglucosammina, e con un residuo di galattosio. Secondo la posizione dell'ultimo legame, le molecole della serie «latto-» sono ulteriormente classificate come tipo 1 (*in basso a sinistra*) e tipo 2 (*in basso a destra*). I glicosfingolipidi «latto-» possono allungarsi e ramificarsi e danno origine a una varietà di molecole che includono gli antigeni dei gruppi sanguigni del sistema *ABH*. I modelli mostrano la conformazione molecolare con il minimo di energia e sono stati costruiti da Steven B. Levery del laboratorio dell'autore e da Ronald E. Stenkamp e Keith D. Watenpaugh dell'Università di Washington.

La membrana plasmatica, l'involucro esterno della cellula, è schematizzata come un doppio strato costituito in primo luogo da molecole fosfolipidiche. I glicosfingolipidi, che sono raffigurati come filamenti di sferette colorate che rappresentano i vari glicidi, sono inclusi nello strato più esterno. I glicidi, idrofili, si trovano alla superficie esterna della membrana nell'ambiente acquoso che circonda la cellula e sono perpendicolari alla regione della molecola che contiene il cerammide idrofobo, il quale serve a fissarla al doppio strato. Molte altre molecole sono incluse nella membrana: tra di esse le glicoproteine, altri tipi di glicolipidi e il colesterolo (non mostrato nell'illustrazione). La parte periferica della catena di carboidrati nei glicosfingolipidi può avere la medesima struttura della parte periferica della catena di carboidrati attaccata alle glicoproteine; le somiglianze sono indicate dalla stessa colorazione. Lo schema a sinistra mostra gli effetti delle interazioni dei glicosfingolipidi e di altre molecole sulla superficie della cellula. I glicosfingolipidi, le cui catene di carboidrati sono esposte alla superficie della membrana cellulare, possono fungere da recettori per i ligandi che si fissano alla cellula (*A*). Essi possono diventare così strettamente associati ad altri glicosfingolipidi (*B*) o a proteine di membrana (*C*). Negli ultimi due casi parte della catena può diventare inaccessibile a potenziali ligandi: allora si dice che il glicosfingolipide è criptico.

Data la diffusione dei glicosfingolipidi nelle cellule e l'importanza delle loro funzioni regolatrici, non sorprende che essi siano coinvolti in numerose gravi malattie. Si pensa che essi siano siti di infezione sulla cellula ospite per diversi tipi di virus e di batteri. Il loro ruolo nel cancro è idoneamente illustrato dal caso della donna con il raro gruppo sanguigno *p*: i tipi di antigeni dei gruppi sanguigni espressi sulla superficie della cellula normale e le relative popolazioni di ogni tipo di antigene si modificano nella cellula con lo sviluppo di tumori maligni. Per apprezzare questo risultato fondamentale si deve però capire che il termine «antigene del gruppo sanguigno» è fuorviante. Questo tipo di antigene si trova non soltanto nel sangue, ma anche in molti altri tipi di tessuto; antigeni per i gruppi sanguigni si trovano in elevate concentrazioni sulla superficie di tutti i tipi di cellule epiteliali che costituiscono il rivestimento mucoso di molti organi. Più del 90 per cento delle neoplasie umane deriva da cellule epiteliali.

La recente identificazione chimica degli antigeni dei gruppi sanguigni come glicosfingolipidi collega l'efficace metodologia immunologica con una imponente massa di informazioni biochimiche. Un'importante conseguenza di questo collegamento consiste nel rendersi conto che i cambiamenti a carico degli antigeni dei gruppi sanguigni nel cancro sono generalizzabili. I biochimici hanno identificato diversi altri glicosfingolipidi che non sono in relazione con antigeni dei gruppi sanguigni, la cui composizione e il cui metabolismo si alterano in modo marcato quando una cellula normale viene trasformata in una cellula neoplastica da vari tipi di virus tumorali e da cancerogeni chimici. L'analisi immunochimica dei glicosfingolipidi nei tumori ha dimostrato che molti antigeni associati ai tumori sono forme chimiche modificate degli antigeni dei gruppi sanguigni.

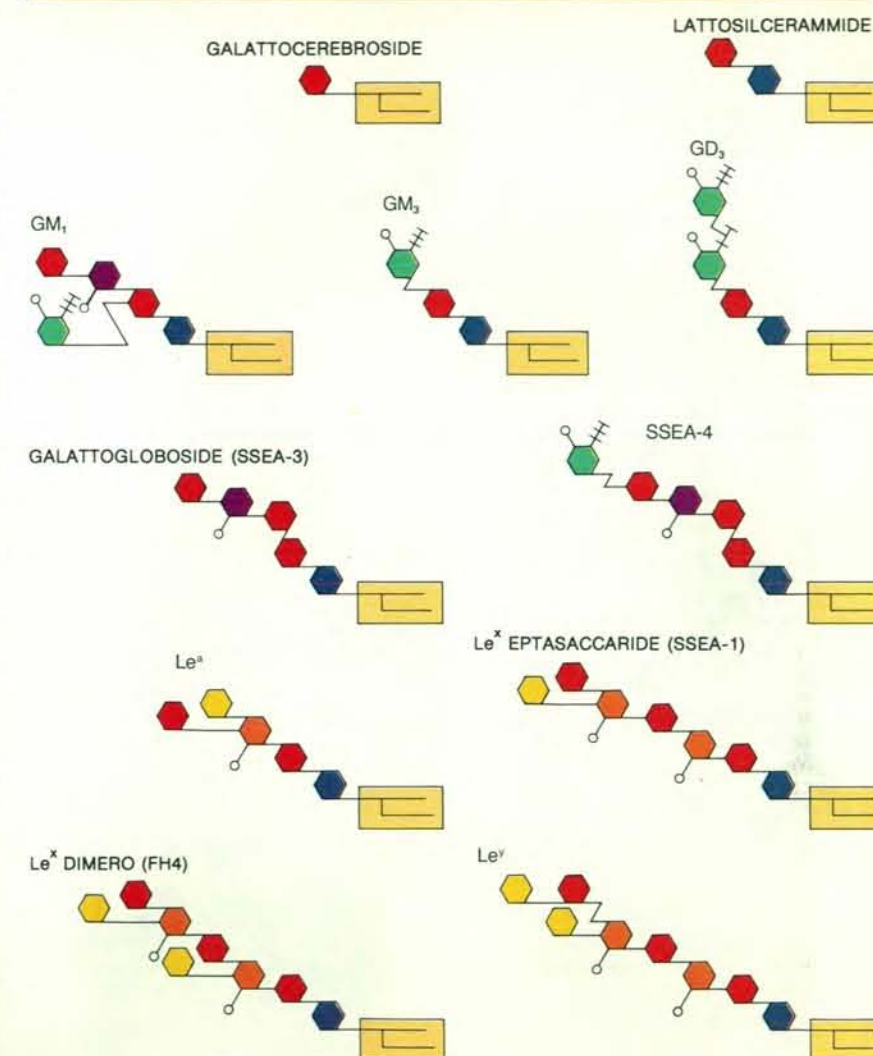
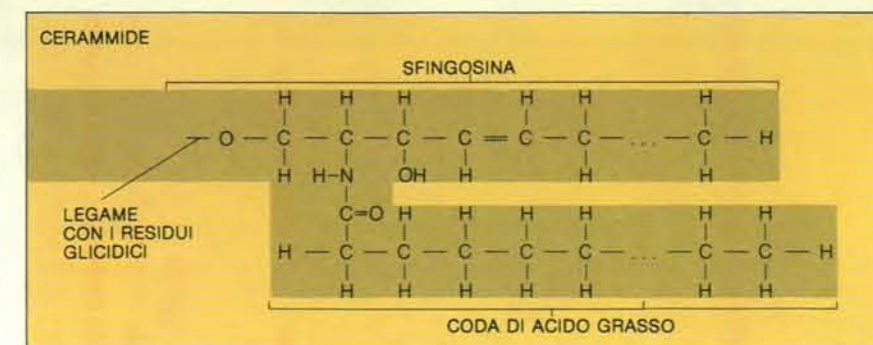
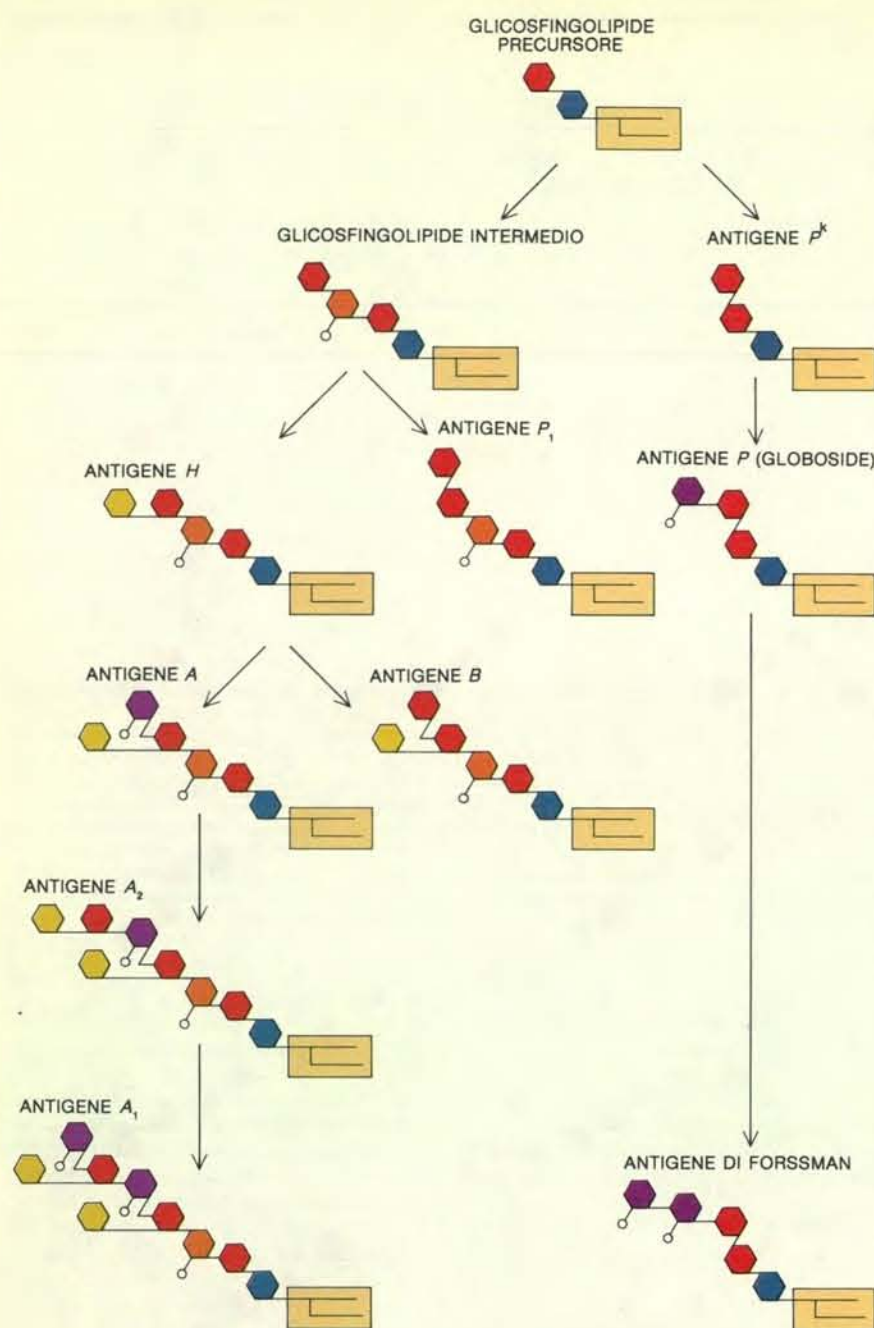
Questi importanti risultati sperimentali hanno suscitato un rinnovato interesse nei riguardi della struttura chimica fondamentale dei glicosfingolipidi e della loro importanza nella vita della cellula. Il primo composto di questo tipo fu scoperto nel 1874 da Johann Ludwig W. Thudichum nel tessuto cerebrale. Thudichum lo chiamò cerebroside. Di recente altri ricercatori hanno trovato che il tessuto cerebrale e quello nervoso sono ricchi di glicosfingolipidi; in effetti, per molti anni lo studio più intensivo di questi composti è stato realizzato da neurobiologi e neurochimici. Tuttavia, il romantico infisso *-sfingo-*, introdotto da Thudichum nel nome che ha attribuito a questi composti, è ancora appropriato. Nella mitologia greca, la Sfinge, il terribile mostro metà donna e metà leone alato, terrorizzava la città di Tebe divorando tutti coloro che non riuscivano a dare una risposta esatta ai suoi enigmi. La funzione di molti glicosfingolipidi nelle strutture neuroniche è ancora un enigma.

La struttura molecolare dei glicosfingolipidi è nota da tempo. Ogni molecola è conformata in un modo che ricorda la lettera L, il cui tratto verticale è costituito da una catena di carboidrati, precisamente di residui glicidici semplici, mentre il tratto orizzontale è un lipide chiamato ceramide (*si veda l'illustrazione a destra*). I due tratti sono così costituiti da componenti diversi: la catena di carboidrati è idrofila e interagisce con l'ambiente acquoso che circonda la cellula; il cerammide, liposolubile, è idrofobo e tende così a escludere le molecole d'acqua presenti nelle vicinanze.

Vi è una considerevole variazione di struttura sia nel cerammide sia nella catena di carboidrati e sono note oggi circa 130 varietà di glicosfingolipidi. Circa 40 sono classificate nella cosiddetta struttura «ganglio-», 10 nella struttura «globo-» e 60 nella struttura «latto-», e ciò dipende dalla sequenza di residui glicidici che si trovano nel cuore della catena di carboidrati e dalla natura dei legami chimici tra glicidi (*si veda l'illustrazione a pagina 19*). I restanti 20 glicosfingolipidi non sono classificati così. Il numero di modi in cui possono essere combinati i componenti di un glicosfingolipide è limitato perché certe strutture carboidratiche sono legate preferenzialmente a certi ceramidi, a seconda delle cellule e dei tessuti in cui esse si trovano.

La stessa sequenza di glicidi che si trova nella struttura «latto-» (ma non in quelle «ganglio-» e «globo-») si trova spesso legata a proteine. Dato che la suddetta sequenza, e non il substrato di ceramide o di proteina a cui i glicidi sono legati, determina spesso la specificità antigenica, gli anticorpi specifici per gli antigeni dei gruppi sanguigni possono anche riconoscere certe glicoproteine. Il fatto che vengano riconosciute le catene glicidiche nelle glicoproteine spiega perché sui rivestimenti mucosi e nelle secrezioni siano presenti gli antigeni dei gruppi sanguigni.

I gangliosidi, che furono scoperti nel 1936 da Ernst Klenk, costituiscono una classe importante di glicosfingolipidi e illustrano efficacemente la distribuzione dei glicosfingolipidi in funzione del tessuto. Klenk li ha trovati solo nel tessuto cerebrale, mentre altri li hanno poi riconosciuti in tutti gli altri tessuti, come componenti delle membrane cellulari. I gangliosidi vengono definiti dalla presenza, nelle loro catene glicidiche, di un particolare acido: l'acido sialico. Certe membrane di cellule nervose sono particolarmente ricche di gangliosidi, le cui catene glicidiche includono diversi tipi di acido sialico. Come Tamio Yamakawa ha suggerito nel 1952, gli schemi di distribuzione variano da cellula a cellula e da specie a specie.

Da un punto di vista energetico le molecole di glicosfingolipidi traggono un notevole vantaggio dal fatto di essere incluse nelle membrane cellulari; la maggior parte dei glicosfingolipidi si trova, infatti, in questa sede. La membrana cellulare viene rappresentata (in un modello semplificato) co-

Nello schema in cui è illustrata la struttura molecolare di importanti glicosfingolipidi, il ceramide è raffigurato in alto; più sotto, compare sempre come un piccolo rettangolo attaccato alle catene di carboidrati. Questi sono rappresentati da esagoni in colore, identificabili con la legenda a sinistra. I glicosfingolipidi nella fila in alto non hanno alcuna delle quattro strutture centrali fondamentali, comuni alla maggior parte di glicosfingolipidi. Le molecole della serie «ganglio-» compaiono nella seconda fila, quelle della serie «globo-» nella terza e quelle della serie «latto-» nelle due file in basso. Il glicosfingolipide Le^a è l'unica molecola raffigurata della serie «latto-» di tipo 1; il resto è costituito da molecole di serie «latto-» di tipo 2.

Gli antigeni dei gruppi sanguigni, di primaria importanza sotto il profilo clinico nella compatibilità con il donatore di sangue e di tessuti, sono glicosfingolipidi. Lo schema mostra le interrelazioni chimiche e strutturali di due sistemi di gruppi sanguigni, rappresentati secondo un codice a colori, che compare nella legenda della pagina precedente. Il sistema meglio noto di gruppi sanguigni è l'*ABH*, ma tutti gli antigeni appartenenti a esso hanno la stessa molecola precursore del sistema di gruppi sanguigni *P* (*in alto*). Tutte le cellule dei tessuti possono essere classificate all'interno di ciascun sistema. Nel sistema *ABH*, le cellule di tipo 0 presentano in superficie una molecola precursore intermedia e l'antigene *H*. Inoltre le cellule di tipo *A* e di tipo *B* mostrano rispettivamente gli antigeni *A* e *B*. Le cellule di tipo *AB*, infine, hanno tutti e quattro gli antigeni. In alcuni soggetti di tipo *A*, detti di tipo A_1, l'antigene *A* può essere allungato e formare l'antigene A_1. In soggetti dell'altro tipo *A*, detti di tipo A_2, l'allungamento della catena è insufficiente a formare l'antigene A_1 e pertanto forma in sua vece l'antigene A_2. Le strutture dei due antigeni *A* sono state determinate da Henrik Clausen e Levery del laboratorio dell'autore. Nel sistema di gruppi sanguigni *t*, le cellule del raro tipo P^k esprimono solo l'antigene P^k e il suo glicosfingolipide precursore. Le cellule di tipo P_2 presentano sia l'antigene P^k sia l'antigene *P*, o globoside, che è derivato dall'antigene P^k. Le cellule di tipo P_1 mostrano l'antigene P_1 oltre gli antigeni del sistema *P*, esposti sulle cellule di tipo P_2. La struttura del sistema *P* è stata stabilita da Donald M. Marcus del Baylor College of Medicine. Circa una persona su cinque porta nei tessuti l'antigene di Forssman, che è derivato dal globoside.

me un doppio strato che consta principalmente di fosfolipidi e di colesterolo. La molecola di fosfolipide, come quella di un glicosfingolipide, ha una testa idrofila e due code idrocarburiche idrofobe. In un contenitore pieno d'acqua le molecole dei fosfolipidi e quelle di colesterolo formano spontaneamente una vescicola sferica, o compartimento a forma di bolla, la cui parete ha lo spessore di due strati di molecole. Nello strato interno, le teste idrofile delle molecole si trovano di fronte l'acqua contenuta nella vescicola, mentre le teste idrofile delle molecole che fanno parte dello strato esterno si trovano di fronte l'acqua al di fuori della vescicola. Le code idrofobe delle molecole di entrambi gli strati della vescicola vengono pertanto mantenute fuori dal contatto con l'acqua.

Le molecole di fosfolipidi sono caratterizzate, nelle membrane cellulari, dalla medesima disposizione generale, ma molte altre sostanze, in particolare le molecole proteiche e i glicosfingolipidi, possono essere incluse nella matrice costituita da fosfolipidi e da colesterolo. In generale, la membrana plasmatica che costituisce la superficie cellulare è molto più ricca di glicosfingolipidi delle membrane degli organelli intracellulari; nella membrana plasmatica poi i glicosfingolipidi sembrano limitati alla parte esterna del doppio strato. Le due code idrocarburiche del cerammide sono inserite nell'interno idrofobo della membrana e conferiscono rigidità strutturale alla membrana stessa. La catena di carboidrati si trova all'esterno della superficie cellulare ed è grosso modo perpendicolare alle code idrocarburiche. Pertanto è esposta alle sostanze extracellulari, a meno che non sia parzialmente mascherata da proteine e da altri glicosfingolipidi che possono essere inclusi nella membrana.

Una delle conseguenze più importanti dell'identificazione chimica degli antigeni dei gruppi sanguigni come glicosfingolipidi è l'aver realizzato che tali antigeni devono essere prodotti genici secondari. In altre parole, la loro struttura non è codificata nel DNA nel modo in cui lo è la sequenza degli amminoacidi in una proteina. Al contrario, un glicosfingolipide viene sintetizzato in reazioni catalizzate da proteine enzimatiche, le glicosiltransferasi.

Le glicosiltransferasi determinano, nella membrana, la sequenza dei residui glicidici: esse riconoscono infatti i nucleotidi, o molecole donatrici di zuccheri, come pure il substrato precursore su cui deve essere trasferito il residuo glicidico proveniente dal nucleotide. Le uniche molecole che prendono parte al processo e le cui strutture sono trascritte direttamente dal codice del DNA sono le proteine enzimatiche note come glicosiltransferasi. Il meccanismo è stato proposto all'inizio per la sintesi dei gangliosidi da Saul Roseman della Johns Hopkins University.

La sintesi dei glicosfingolipidi potrebbe essere controllata in due modi. Il primo è comune alla sintesi di molti altri enzimi: la trascrizione della sequenza delle basi del DNA e la loro traduzione in proteine enzimatiche può essere accelerata o inibita dall'attivazione o inibizione di speciali sequenze di DNA, chiamate promotori o potenziatori. In alternativa, l'attività sintetica delle glicosiltransferasi nella membrana potrebbe venire alterata da modificazioni, relativamente minori, delle molecole enzimatiche stesse dopo che sono state sintetizzate. La modificazione chimica degli enzimi può influenzarne la distribuzione all'interno della membrana e le interazioni con i substrati in seno a essa. Inoltre, l'espressione dei glicosfingolipidi sulla superficie delle membrane dipende non soltanto dalla velocità della loro sintesi, ma anche dal modo in cui essi e altre molecole di membrana sono riuniti sulla membrana stessa. La loro accessibilità agli anticorpi o a sostanze esterne può essere influenzata dalla vicinanza ad altre molecole (*si veda l'illustrazione a pagina 20*).

Dato che l'espressione dei glicosfingolipidi è controllata da più di un unico meccanismo, la concentrazione di questi composti alla superficie delle cellule può essere molto più sensibile a impercettibili variazioni ambientali della concentrazione, per esempio, di una proteina. Rapidi cambiamenti nell'espressione dei glicosfingolipidi potrebbero venir determinati da una regolazione coordinata dei geni che codificano per gli enzimi che sono indispensabili alla sintesi dei glicosfingolipidi, da variazioni nella competizione che si stabilisce tra certi enzimi per il substrato disponibile, oppure da un cambiamento nell'orientazione e nell'esposizione di antigeni di natura glicosfingolipidica già presenti sulla superficie cellulare. Ricercatori nel mio e in altri laboratori hanno stabilito che certi glicosfingolipidi si trovano su tipi specifici di cellule per brevi permanenze nel corso dello sviluppo embrionale. Queste risultanze fanno pensare che essi siano collegati a meccanismi di crescita e di differenziamento cellulari. Può darsi che le cellule trovino più redditizio modificare i glicosfingolipidi che modificare le proteine come reazione alle mutevoli circostanze dello sviluppo.

La comparsa e la scomparsa degli antigeni glicosfingolipidici, in funzione dello stadio di sviluppo, sono raffigurate sotto forma di grafico per un embrione di topo in una fase che precede l'impianto. L'antigene *SSEA-1* (l'acronimo sta per *stage-specific embryonic antigen*, antigene embrionale stadio-specifico) compare tra lo stadio di 8 e quello di 32 cellule e declina marcatamente dopo che le cellule si riuniscono e formano un ammasso compatto, in uno stadio detto appunto di compattamento (*curva in colore*). Quando *SSEA-1* diminuisce, compare l'antigene Le^y (*curva in grigio*); questo è chimicamente molto simile al primo, nel senso che in corrispondenza del galattosio terminale vi è un ulteriore residuo di fucosio. Gli antigeni *SSEA-3* e *SSEA-4* continuano a essere espressi con intensità fino allo stadio precoce del compattamento e scompaiono quasi completamente a partire dallo stadio a 32 cellule (*curva in nero*). Questi glicosfingolipidi possono essere strettamente associati con il meccanismo di regolazione del riconoscimento tra le varie cellule e la crescita dei tessuti.

Sembra dunque evidente che i glicosfingolipidi possono influire sulla cellula e sulla sua crescita in due modi principali. Possono modulare le funzioni delle proteine presenti nella membrana plasmatica e servono, assieme a queste proteine, da marcatori di superficie, necessari per mantenere una confacente comunicazione intercellulare. I primi indizi che i glicosfingolipidi possano influenzare il funzionamento delle proteine di membrana sono emersi dal fatto che i glicosfingolipidi e le proteine sono strettamente associati all'interno della membrana. Questa prossimità è stata dimostrata da Clifford Lingwood del mio laboratorio e da Tae Hwa Ji dell'Università del Wyoming. Ranwel Caputto e collaboratori dell'Università di Cordoba in Argentina hanno quindi trovato che i glicosfingolipidi nel cervello possono attivare l'ATPasi, una proteina enzimatica che libera l'energia necessaria alla trasmissione di messaggi lungo la fibra nervosa.

Le proteine recettrici (o recettori proteici) incluse nella membrana plasmatica fungono spesso da legame essenziale nella complessa sequenza di eventi che conduce alla mitosi, o divisione di cellule somatiche. Recenti studi di Eric Bremer, del mio laboratorio, con Daniel F. Bowen-Pope, Elaine W. Raines e Russell Ross, dell'Università di Washington, indicano che certi glicosfingolipidi possono rompere tale legame e quindi inibire la crescita cellulare.

Le cellule animali vanno generalmente incontro alla mitosi e proliferano solo in presenza di ormoni specifici, extracellulari, i fattori di crescita. Questi si legano a specifiche proteine recettrici in corrispondenza di siti sul lato esterno della membrana plasmatica. La proteina recettrice è una grossa molecola contenuta nella membrana, in parte all'interno e in parte all'esterno della cellula. La sua componente citoplasmatica è un enzima che catalizza - in seno al citoplasma - l'attacco di gruppi fosfato a varie proteine, tra cui lo stesso recettore. Il legame del fattore di crescita può condurre all'attivazione dell'enzima associato con la componente citoplasmatica del recettore e induce l'attacco al recettore di un gruppo fosfato. Questo attacco sembra provocare un cambiamento conformazionale nel recettore e sono proprio questi tipi di cambiamento a far sì che i recettori si aggreghino sulla superficie della membrana. I recettori con attaccati i rispettivi fattori di crescita sono quindi assunti all'interno della cellula.

I ricercatori hanno trovato che, quando i glicosfingolipidi GM_1 o GM_3 vengono aggiunti a cellule che proliferano in presenza di fattori di crescita, la proliferazione cellulare viene inibita. Inoltre, nelle cellule ottenute da preparati di membrane e la cui proliferazione è stata inibita da GM_1 o da GM_3, le proteine recettrici non si attaccano facilmente a un gruppo fosfato. L'effetto sembra essere specifico per GM_1 e GM_3; nessun altro glicosfingolipide ha portato agli stessi risultati. Abbiamo pertanto ritenuto che il funzionamento del recettore fosse sensibile all'ambiente creato dai glicosfingolipidi nelle membrane. È degno di menzione il fatto che, quando le cellule vengono trasformate da virus tumorigeni, le risultanti cellule tumorali presentano un contenuto molto più basso di GM_3 in alcuni tipi di tumore e di GM_1 in altri tipi di tumore. Queste riduzioni nella concentrazione possono essere correlate con una perdita del controllo sulla crescita delle cellule esercitato dai glicosfingolipidi.

I glicosfingolipidi però non sono sempre associati all'inibizione della crescita e del differenziamento della cellula. Studi di Yoshitaka Nagai e collaboratori dell'Università di Tokyo indicano che l'aggiunta di un altro glicosfingolipide, chiamato GQ_{1b}, a cellule nervose embrionali tumorali indu-

ce la formazione di cellule nervose mature e molti altri cambiamenti. Più di recente Masaki Saito, Hisao Nojiri e i loro collaboratori dell'Università Jichi in Giappone hanno osservato che, quando certe cellule leucemiche di topo sono incubate con GM_3, si differenziano in macrofagi con un più elevato contenuto di GM_3.

Oltre al ruolo che esercitano nella regolazione delle proteine, i glicosfingolipidi svolgono funzioni cellulari di propria competenza. La loro funzione più importante consiste nel codificare a livello cellulare differenze tra specie, tra individui della stessa specie e perfino tra cellule di un dato individuo. Gli antigeni dei gruppi sanguigni, per esempio, mostrano in che modo i glicosfingolipidi possono variare tra i singoli individui della popolazione umana.

Recentemente è aumentato l'interesse per i cambiamenti che subiscono, nello sviluppo normale come pure nella crescita neoplastica, i glicosfingolipidi marcatori specifici per ogni tipo di cellula. Per esempio, Davor Solter e Barbara B. Knowles del Wistar Institute of Anatomy and Biology di Philadelphia hanno stabilito l'esistenza di un antigene chiamato *SSEA-1*. (L'acronimo significa antigene embrionale stadio-specifico, dall'inglese *stage-specific embryonic antigen*.) *SSEA-1* non reagisce con il proprio anticorpo nell'uovo fecondato, ma può essere individuato in qualche momento tra la terza e la quinta divisione cellulare, cioè in embrioni costituiti da otto a 32 cellule. A questo punto le cellule subiscono un processo di consolidamento e cominciano ad aderire reciprocamente in modo da sviluppare i contatti intercellulari. Completata questa fase, la concentrazione di *SSEA-1* declina rapidamente.

La struttura chimica di *SSEA-1* è stata identificata da Reiji Kannagi, Steven B. Levery ed Edward Nudelman del mio laboratorio e da Ten Feizi e collaboratori del Clinical Research Centre del Medical Research Council di Harrow, in Inghilterra, come una catena di carboidrati, di nome Le^x, su un glicosfingolipide o una glicoproteina. Bruce Fenderson con Uri Zehavi del mio laboratorio e Susan J. Kimber dell'MRC Laboratory for Molecular Biology di Carshalton, in Inghilterra, hanno trovato che la catena Le^x, o la sua struttura coniugata, funge da inibitore del consolidamento nello stadio embrionale che segna il passaggio da 16 a 32 cellule. Presumibilmente la struttura Le^x su una cellula avverte la presenza di altre cellule vicine e interferisce poi con il complesso processo di adesione cellulare, che potrebbe essere basato su una affinità specifica tra la catena Le^x sulla superficie cellulare e il suo recettore. Il consolidamento può avvenire tempestivamente con il declino della catena dopo lo stadio a 32 cellule.

Sembra che molte tossine batteriche e organismi virali e microbici siano in grado di sfruttare la capacità dei glicosfingolipidi di mediare le interazioni della cellula con l'ambiente. Una trentina di anni fa, William E. van Heyningen dell'Università di Oxford ha riportato il primo caso di gangliosidi che interagiscono con la tossina tetanica. Rilievi successivi suggeriscono che il ganglioside GM_1 sia il recettore specifico per tale tossina e per quella del colera. La tossina colerica agisce non invadendo la cellula, ma provocando da parte della membrana una perdita di impermeabilità all'acqua e agli elettroliti idrosolubili, presenti all'interno della cellula. Dopo van Heyningen si è cercato di sapere se altri glicosfingolipidi siano recettori di altre tossine e fattori bioattivi. I ricercatori hanno dimostrato che i gangliosidi interagiscono - anche se non necessariamente come recettori - con molte sostanze biologicamente importanti, tra le quali la tossina botulinica, l'interferon, l'interleuchina, la serotonina, ormoni e virus Sendai.

Parecchi gruppi di ricercatori hanno trovato di recente che glicosfingolipidi specifici, presenti sulla superficie di una cellula ospite, interagiscono con le proteine dei parassiti virali e batterici. In Finlandia e in Svezia è stato scoperto che certi ceppi di *Escherichia coli*, batteri che provocano un'infezione delle vie urinarie e dei reni, la pielonefrite, presentano fini strutture fibrillari, le fimbrie, sulla loro superficie cellulare. Le proteine di queste fimbrie interagiscono con un glicosfingolipide, il globoside, e i suoi precursori chimici; questa interazione è probabilmente la prima fase nell'infezione della cellula ospite.

Di recente, Karl-Anders Karlsson e collaboratori dell'Università di Göteborg hanno marcato cellule di *E. coli*, che causano infezioni urinarie, con un marcatore radioattivo; hanno poi osservato le interazioni batteriche con i glicosfingolipidi che erano stati separati su un sottile strato di silicagel. Dei 32 glicosfingolipidi saggiati gli unici che si sono attaccati ai batteri erano quelli che possedevano una particolare catena di tre glicidi legata alla componente cerammidica della molecola. Catene di altri glicidi, attaccate alla catena iniziale di tre glicidi, non avevano alcun effetto sull'unione con i batteri. Le proteine di superficie dei batteri dovevano aver dunque appreso a riconoscere corte catene interne di glicidi, che potevano essere esibite sulla cellula ospite da vari glicosfingolipidi.

Il riconoscimento tra cellule è illustrato nel modo che si pensa abbia luogo tra due cellule animali e tra un batterio e una cellula ospite. Le proteine di membrana su ogni cellula animale (*a sinistra*) possono riconoscere la struttura di un antigene *SSEA-1*, esposto sulla superficie dell'altra cellula. Un riconoscimento del genere è tipico in embrioni che hanno da otto a 32 cellule, prima dell'impianto. Le proteine delle fimbrie in un batterio *Escherichia coli* patogeno (*a destra*) possono riconoscere i glicidi interni di glicosfingolipidi della serie «globo-» in una cellula ospite e utilizzarli come siti di infezione. I glicidi dei glicosfingolipidi sono colorati secondo un codice, per il quale si rimanda alla legenda di pagina 21.

Le tecniche di Karlsson hanno anche permesso di determinare i siti di legame di certi tipi di batteri che costituiscono la normale «flora intestinale». Batteri come questi sono innocui e arrecano beneficio all'organismo che li ospita inibendo lo sviluppo di batteri patogeni. Per esempio, il comune batterio intestinale chiamato *Propionibacterium* ha una proteina di superficie che si lega a un glicosfingolipide molto comune, il lattosilcerammide. La struttura carboidratica di questo composto è il precursore della maggior parte dei glicosfingolipidi noti. Infatti i due glicidi attaccati alla componente cerammidica del lattosilcerammide sono identici ai primi due glicidi della catena di tre, riconosciuta dai batteri della specie *Escherichia coli* che provocano infezioni urinarie.

Mentre le tossine e le infezioni batteriche e virali possono trarre vantaggio da glicosfingolipidi esistenti su una cellula sana, la crescita neoplastica è chiaramente associata a glicosfingolipidi alterati. Senza dubbio questa alterazione esprime una qualche anomalia che ne è alla base. Qualunque ne sia la causa, uno scompiglio nelle popolazioni di glicosfingolipidi potrebbe interferire con una parte essenziale delle risorse della cellula per il mantenimento di una ordinata «vita sociale». Non sarebbe sorprendente trovare che le alterazioni dei glicosfingolipidi siano una causa delle interazioni sociali caotiche e indisciplinate, caratteristiche delle cellule neoplastiche.

Il tipo di alterazione che si realizza nei glicosfingolipidi dipende dalla cellula ospite oltre che dall'agente cancerogeno specifico. Alcune cellule tumorali accumulano glicosfingolipidi semplici perché la sintesi dei glicosfingolipidi più complessi ai quali danno origine quelli semplici è in una certa misura bloccata. Altre cellule tumorali, particolarmente le cellule di carcinoma epiteliale umano, sintetizzano catene insolite di carboidrati e pertanto accumulano nuovi glicosfingolipidi: i neoglicolipidi.

I glicosfingolipidi e i neoglicolipidi precursori sono stati entrambi originariamente identificati come antigeni associati a tumori negli studi convenzionali con gli anticorpi di coniglio. Tuttavia la consapevolezza che le cellule tumorali possono generare glicosfingolipidi varianti non ha potuto essere sfruttata appieno fino a quando, nel 1976, George Köhler e Cesar Milstein del Medical Research Council Laboratory di Cambridge hanno messo a punto la tecnica per produrre anticorpi monoclonali. Questi anticorpi hanno in comune un'affinità estremamente specifica per un singolo antigene e pertanto la presenza di un dato antigene può essere rapidamente rilevata attaccando all'anticorpo monoclonale una molecola marcatore. Per ottenere elevate quantità di anticorpi monoclonali, specifici per gli antigeni tumorali, si devono in primo luogo iniettare questi antigeni nei topi. Cellule di milza di topi immunizzati vengono quindi fuse con cellule tumorali di topo. Le cellule risultanti, gli ibridomi, combinano la capacità, propria delle cellule tumorali, di moltiplicarsi illimitatamente con la capacità, propria delle cellule della milza, di generare anticorpi. In tutto sono stati identificati e caratterizzati sotto il profilo chimico oltre 10 tipi di antigeni di natura glicosfingolipidica, associati a tumori e definiti da anticorpi monoclonali specifici.

Poiché differenti ibridomi danno origine ad anticorpi diversi, essi devono in qualche modo essere selezionati. La maggior parte degli esperti di immunologia dei tumori impiega un procedimento casuale: inietta antigeni indeterminati di tessuti tumorali in topi e quindi seleziona i risultanti ibridomi in base alla loro capacità di reagire con gli antigeni associati con i tumori. Nel nostro laboratorio adottiamo un metodo meno convenzionale, per mezzo del quale la selezione viene effettuata, in effetti, all'inizio. In primo luogo otteniamo e caratterizziamo chimicamente un glicosfingolipide associato con un tipo specifico di tumore. Il glicosfingolipide viene quindi posto a rivestimento dei batteri e questi vengono inoculati in topi. Nella selezione possiamo quindi ricercare il tipo di ibridoma che reagisce in modo preferenziale con l'antigene di natura glicosfingolipidica che era stato preparato in origine.

Abbiamo utilizzato la nostra procedura per ottenere anticorpi monoclonali diretti contro una varietà di nuove strutture associate ai tumori. Nei loro studi sugli antigeni Le^x e Le^y, regolati dallo sviluppo, Nudelman, Levery e Kannagi hanno trovato che molti tessuti tumorali contengono una notevole quantità di glicosfingolipidi strettamente affini a Le^x e Le^y. I nuovi antigeni includono le forme dimere e tri-

L'antigene *FH4*, associato ai tumori, viene rappresentato da questo modello generato per mezzo di una simulazione al calcolatore. L'area in grigio, nello schema riportato a sinistra, è una regione idrofoba che può interagire con l'anticorpo. L'antigene *FH4* è stato identificato come il glicosfingolipide Le^x dimero e si esprime in modo molto intenso nei tumori umani che derivano da cellule gastrointestinali. Anche questo modello è stato realizzato da Levery, del laboratorio dell'autore, assieme a Stenkamp e Watenpaugh dell'Università di Washington.

mere di Le^x e forme alle quali è attaccato l'acido sialico. Yasuo Fukushi e Kazuo Abe nel mio laboratorio hanno quindi prodotto con successo anticorpi monoclonali che sono in grado di distinguere tra loro i diversi tipi di Le^x.

Uno di questi anticorpi, denominato anticorpo *FH4*, reagisce in maniera specifica con l'antigene Le^x dimero, associato ai tumori, ma non con la forma monomera. Fukushi ha quindi saggiato la capacità dell'anticorpo *FH4* marcato di formare un legame con i tessuti sani di embrioni, ha colorato il complesso, evidenziando così l'intensa espressione del Le^x dimero nelle cellule epiteliali dello stomaco di embrioni dell'età di 35-45 giorni. La colorazione era tuttavia limitata agli epiteli gastrico e intestinale di feti di 100 giorni, caratterizzati da cripte profonde. In pratica scompare nelle cellule del neonato e dell'adulto (*si veda l'illustrazione qui sotto*).

Gli antigeni della superficie cellulare che hanno questo comportamento (antigeni oncofetali), si trovano sulle cellule tumorali con la stessa abbondanza relativa con cui si trovano in brevi stadi dello sviluppo di certi tipi di cellule embrionali. Oggi ne sono noti diversi. Un altro è, per esempio, un derivato del Le^x dimero, contenente acido sialico e definito sotto il profilo antigenico dall'anticorpo *FH6*. Colora intensamente i tessuti fetali, non colora il tessuto gastrointestinale adulto mentre colora in maniera intensa e specifica i carcinomi gastrointestinali. Questi risultati confermano l'osservazione che viene fatta di frequente e per la quale le cellule tumorali assomigliano a cellule relativamente indifferenziate, presenti nelle fasi precoci dello sviluppo embrionale.

Un altro importante antigene oncofetale correlato con i gruppi sanguigni è il derivato contenente acido siàlico di un glicosfingolipide, l'antigene Le^a. Un anticorpo monoclonale, diretto contro questo antigene, è stato prodotto da Hilary Koprowski e collaboratori al Wistar Institute; l'antigene è stato in seguito isolato e caratterizzato sotto il profilo chimico da Victor Ginsburg, John L. Magnani e collaboratori al National Institute of Arthritis, Diabetes and Digestive and Kidney Diseases. Dato che gli anticorpi monoclonali si legano solo alla catena di carboidrati di un glicosfingolipide possono anche riconoscere queste catene su glicoproteine del siero non attaccate a cellule. Per esempio, gli anticorpi contro i derivati contenenti acido sialico dei glicosfingolipidi Le^a e Le^x reagiscono con i sieri dei pazienti affetti da cancro. Questi antigeni circolanti possono essere facilmente saggiati e pertanto la loro presenza è utile per la diagnosi dei tumori. Il loro studio è un capitolo importante delle indagini cliniche.

Comunque, gli antigeni glicosfingolipidici attaccati alle cellule rappresentano probabilmente la miglior promessa per la terapia del cancro. Per esempio, il normale Le^x e i suoi analoghi Le^x e Le^y dimeri e trimeri sono attaccati alle cellule tumorali e non liberi nel siero. In teoria, gli anticorpi specifici per gli antigeni associati a tumori, quando vengono identificati, possono essere prodotti e somministrati a malati di cancro. Le cellule neoplastiche che presentano in superficie popolazioni alterate di antigeni associati a tumori potrebbero essere allora ricercate dagli anticorpi e distrutte con meccanismi scatenati dagli anticorpi stessi e da certi tipi di cellule «killer». La graduale scomparsa del tumore maligno nella donna con il raro gruppo sanguigno *p* si era realizzata inavvertitamente proprio in questo modo.

L'espressione dell'antigene *FH4* è illustrata nello sviluppo umano normale e in un tumore gastrointestinale. Le colorazioni in bruno intenso, che si notano nelle microfotografie, sono dovute al fatto che l'anticorpo monoclonale *FH4* si è legato con il rispettivo antigene. Le colorazioni della microfotografia in alto a sinistra rivelano un'espressione molto intensa dell'antigene nelle cellule epiteliali gastriche di un embrione di età poco superiore alle cinque settimane (38 giorni). La microfotografia in alto a destra mostra le cellule epiteliali gastriche di un feto di 120 giorni; soltanto le profonde pieghe della mucosa dello stomaco appaiono intensamente colorate. Nell'epitelio gastrico adulto (*in basso a sinistra*), praticamente non si nota alcuna colorazione dei tessuti. In basso a destra, infine, le cellule tumorali di un carcinoma allo stomaco appaiono colorate, mentre quelle normali no. Gli antigeni espressi alla superficie delle cellule tumorali e nello sviluppo del feto vengono definiti oncofetali.

L'intensità con cui l'antigene *FH4* viene espresso è rappresentata in grafico nel caso dello sviluppo fetale umano. Nelle cellule epiteliali dello stomaco l'espressione di *FH4* raggiunge la massima intensità quando l'embrione si trova tra le cinque e le sette settimane di gestazione. Negli epiteli del colon e del tenue l'intensità massima si osserva in età comprese tra sette e nove settimane. L'antigene *FH4* si esprime in tumori differenziati dello stomaco o del colon proprio come avviene negli epiteli fetali (*A, B*). Nei tumori non differenziati dello stomaco o del colon l'*FH4* non si esprime; queste cellule neoplastiche assomigliano alle cellule dell'embrione dei primi stadi di vita (*C, D*).

Recenti studi di immunologia hanno dimostrato che molti anticorpi monoclonali, diretti contro antigeni associati a tumori, sono diretti in effetti contro i glicosfingolipidi. I risultati di almeno due studi clinici giustificano apparentemente alcune delle più ottimistiche previsioni. In uno dei due studi, un gruppo di ricercatori sotto la guida di Lloyd J. Old e Kenneth O. Lloyd dello Sloan-Kettering Institute for Cancer Research di New York ha prodotto un anticorpo con una elevata affinità per il glicosfingolipide GD_3. Alan N. Houghton e collaboratori allo Sloan-Kettering Institute lo hanno somministrato a 12 pazienti affetti da melanoma, una forma virulenta di cancro che produce sulla cute dei tumori di colore scuro, che alla fine danno luogo a metastasi, invadendo vari organi. In tre dei 12 pazienti è stata osservata una regressione della malattia e quattro altri pazienti hanno avuto una risposta mista.

Nel secondo studio, Ronald B. Herberman del National Cancer Institute ha somministrato un anticorpo anti-GD_3 ad altri 12 pazienti affetti da melanoma. L'anticorpo era stato preparato da Ralph A. Reisfeld e collaboratori della Scripps Clinic and Research Foundation, i quali ne hanno anche stabilito la specificità antigenica. Anch'esso è stato efficace in tre pazienti su 12, anche se altri tipi di anticorpi antimelanoma non hanno mostrato alcun effetto. Inoltre, in seguito al trattamento con anticorpi, nessun paziente ha avuto disturbi a carico di rene, intestino e retina, anche se il glicosfingolipide GD_3 è presente in quantità esigua nei tessuti normali del primo e del secondo organo e in quantità elevate nella retina. Unici effetti nocivi sono state delle reazioni cutanee infiammatorie attorno al melanoma. Questi risultati sono incoraggianti e suggeriscono che gli antigeni glicosfingolipidici nelle cellule tumorali siano più accessibili a un attacco da parte degli anticorpi di quanto lo siano gli stessi antigeni nei tessuti normali.

Sono state proposte parecchie strategie per rimuovere o uccidere cellule tumorali; esse potrebbero sfruttare la specificità degli anticorpi monoclonali. Isotopi radioattivi fissati sull'anticorpo potrebbero legarsi alle cellule tumorali e quindi irradiarle oppure indicare con precisione la loro posizione nel corpo. Seguirebbe o un ulteriore trattamento di irradiazione oppure un intervento chirurgico. Anche farmaci fissati sugli anticorpi potrebbero uccidere selettivamente le cellule tumorali.

Inoltre, è bene non dimenticarlo, gli anticorpi costituiscono gli strumenti naturali ed estremamente efficaci che il corpo impiega per isolare e uccidere le cellule malate. Un anticorpo monoclonale, come qualsiasi altro anticorpo, non solo si lega a una cellula che presenta in superficie un antigene, ma anche la contrassegna in modo che venga distrutta dal sistema immunitario dell'ospite. Molti ricercatori sono oggi impegnati nel tentativo di trasformare qual fortunato caso clinico di 35 anni fa in un trattamento efficace e di routine per molte forme di cancro.

AMBIENTE E INQUINAMENTI

LE SCIENZE *edizione italiana di* SCIENTIFIC AMERICAN

ha dedicato all'argomento diversi articoli:

Il clima di Marte

Anche se un tempo era simile al clima primordiale della Terra, ha subito una diversa evoluzione passando da condizioni miti all'intenso freddo attuale che in inverno provoca il congelamento di anidride carbonica ai poli

di Robert M. Haberle

Nell'emisfero settentrionale di Marte è arrivato l'autunno. Alle medie latitudini la temperatura media scende a oltre 70 gradi Celsius sotto zero; al polo nord ha probabilmente già raggiunto −123 gradi, abbastanza freddo da far congelare l'anidride carbonica, il componente principale della tenue atmosfera marziana. Una calotta di ghiaccio di anidride carbonica si sta formando sul polo e, prima che l'inverno sia finito, avrà raggiunto il cinquantesimo parallelo. Intanto, nell'emisfero meridionale l'inverno è finito e dal polo sud sublima anidride carbonica. Lungo il margine della calotta polare che si ritira, il violento contrasto termico tra il ghiaccio e il suolo riscaldato dal Sole provoca forti venti. Durante la breve, ma calda estate meridionale, allorché Marte si trova alla minima distanza dal Sole, i venti solleveranno dalla superficie tempestosi vortici di polvere. E la polvere potrà così avvolgere l'intero pianeta.

Questi eccezionali particolari delle condizioni climatiche di Marte sono oggi noti grazie alle missioni spaziali degli anni sessanta e settanta, in particolare grazie alla missione Viking. Per quanto aspro e inospitale secondo i parametri terrestri, il clima marziano è per alcuni aspetti simile a quello della Terra. Su entrambi i pianeti, per esempio, domina un sistema globale di venti caratterizzato dalla presenza di alisei ai tropici e di perturbazioni cicloniche alle medie latitudini. Tuttavia le differenze sono ancor più considerevoli. Mentre gli elementi gassosi prevalenti sulla Terra sono azoto e ossigeno, l'atmosfera di Marte è per il 95,3 per cento anidride carbonica. Inoltre, l'atmosfera marziana ha una massa molto inferiore. La pressione da essa esercitata sulla superficie del pianeta è in media di soli 6,1 millibar, rispetto alla media terrestre di 1013 millibar. Nell'attesa che il prossimo veicolo spaziale arrivi su Marte - *Mars Observer*, il cui lancio è in programma per il 1990 - i ricercatori continuano a studiare i dati dei Viking. Nel corso del lavoro essi ottengono una migliore conoscenza di come si è evoluto il clima di Marte, simile al clima della Terra per certi aspetti e radicalmente differente per altri.

All'inizio dei 4,6 miliardi di anni di storia del sistema solare il clima di Marte e quello della Terra erano probabilmente molto più simili di quanto siano attualmente. Sembra infatti probabile che i due pianeti avessero inizialmente atmosfere di composizione approssimativamente analoga. Si ritiene che su entrambi i pianeti l'atmosfera primordiale fosse costituita da gas liberati da masse rocciose o per degassamento vulcanico (un processo in cui i gas vengono riscaldati fino a fuoriuscire dal magma interno al pianeta e quindi espulsi attraverso vulcani o altre bocche) oppure per degassamento da impatto (un processo in cui i gas vengono liberati da collisioni di masse rocciose contro il pianeta in fase di accrescimento). Le rocce stesse si formavano dalla condensazione di gas e polvere della nebulosa solare; la loro composizione era determinata dalla temperatura alla quale si formavano, a sua volta dipendente dalla distanza del pianeta dal Sole. Marte e la Terra si condensarono a distanze abbastanza simili perché la composizione delle loro rocce, e quindi delle loro atmosfere primordiali, sia fondamentalmente analoga.

Vi sono validi motivi per ritenere che le atmosfere primordiali fossero costituite principalmente da vapore acqueo, anidride carbonica e azoto. Sulla base di considerazioni teoriche si ritiene che questi gas dovrebbero essere liberati da magmi caldi e le osservazioni delle emissioni dai vulcani terrestri confermano questa supposizione. Le atmosfere della Terra e di Marte sono oggi così differenti perché le sostanze volatili di partenza hanno avuto sui due pianeti un destino diverso.

La maggior parte dell'acqua terrestre si trova attualmente negli oceani; infatti la quantità espulsa in forma gassosa era molto maggiore di quanta ne potesse contenere l'atmosfera. La presenza degli oceani alimenta a sua volta un ciclo idrologico attivo che sottrae anidride carbonica alla atmosfera: l'acqua dapprima evapora dagli oceani e poi, ricadendo sotto forma di pioggia, porta con sé anidride carbonica. Quando le masse rocciose continentali sono alterate dall'acqua piovana, l'anidride carbonica in essa disciolta reagisce con i silicati formando ioni bicarbonato che, trasportati dai fiumi, si depositano alla fine sul fondo degli oceani come carbonato di calcio (calcare). La maggior parte dell'anidride carbonica espulsa per degassamento è attualmente conservata in questi serbatoi rocciosi; al confronto, la quantità che permane nell'atmosfera è insignificante; l'azoto è invece divenuto il componente dominante dell'atmosfera terrestre (78,1 per cento del volume totale) poiché esso è chimicamente inerte e relativamente insolubile in acqua. L'ossigeno, che costituisce il 20,9 per cento dell'atmosfera terrestre, deve la sua abbondanza unicamente all'attività fotosintetizzante delle piante.

Anche su Marte venne espulsa una quantità d'acqua superiore a quella che poteva esser contenuta dall'atmosfera. Si è stimato che la quantità totale di acqua fuoriuscita per degassamento avrebbe coperto il pianeta con un oceano profondo da decine a centinaia di metri. Oggi Marte è troppo freddo per ospitare oceani o qualsiasi altra forma d'acqua allo stato liquido. Il motivo più evidente è che la sua distanza dal Sole è una volta e mezzo quella della Terra e riceve quindi, rispetto a quest'ultima, meno della metà dell'energia solare.

Il polo sud di Marte è coperto da ghiaccio di anidride carbonica e la calotta è forse perenne. La fotografia è stata ripresa dal modulo orbitale di *Viking 2* nel 1977, durante l'estate meridionale. Al polo nord, invece, tutto il ghiaccio di anidride carbonica sublima ogni estate, mettendo allo scoperto una calotta di ghiaccio d'acqua. L'anidride carbonica costituisce circa il 95 per cento della tenue atmosfera marziana e circa il 20 per cento solidifica ogni inverno al polo sud. L'immagine è stata elaborata dallo US Geological Survey di Flagstaff in Arizona.

Tuttavia la distanza dal Sole non può essere l'unico fattore, poiché sembra che Marte fosse una volta sufficientemente caldo da avere in superficie corsi d'acqua.

Le immagini della superficie del pianeta riprese da *Mariner 9* e dai moduli orbitali dei Viking hanno rilevato numerosi canali il cui aspetto farebbe supporre l'azione erosiva di corsi d'acqua. I canali non provano necessariamente l'esistenza di un clima più caldo, infatti alcuni potrebbero essere stati formati da corsi d'acqua sottostanti a uno strato isolante di ghiaccio superficiale in un clima sostanzialmente non più caldo dell'attuale. Molti dei canali, tuttavia, somigliano a reti di deflusso superficiale, con affluenti convergenti in una singola valle. Supponendo che questi canali siano stati scavati da corsi d'acqua superficiali, allora la temperatura media globale deve essere stata a quel tempo superiore a 0 gradi Celsius (oggi è di circa −53 gradi). Dal numero di crateri meteoritici sovrapposti ai canali di deflusso, si è stimato che l'età di questi ultimi si aggiri intorno ai quattro miliardi di anni. Ne deriva quindi che il clima di Marte è stato relativamente più caldo per il primo mezzo miliardo di anni della sua storia.

Risulta lecito chiedersi come potesse la temperatura su Marte mantenersi al di sopra del punto di congelamento dell'acqua. Il meccanismo più probabile è il ben noto «effetto serra». I gas dell'atmosfera marziana sono pressoché trasparenti alla maggior parte della radiazione solare, e perciò l'atmosfera non è direttamente riscaldata dal Sole. Come l'atmosfera terrestre, essa viene invece riscaldata dal basso, dall'energia reirradiata dalla superficie a lunghezze d'onda più elevate, quelle dell'infrarosso. A queste lunghezze d'onda, infatti, i gas assorbono e catturano il calore. Il vapore acqueo e l'anidride carbonica sono gas particolarmente efficienti nel produrre l'effetto serra.

L'effetto serra può essere misurato confrontando la temperatura alla quale un pianeta irradia l'energia assorbita dal Sole con la sua temperatura superficiale osservata. La temperatura effettiva di irradiazione della Terra è di −18 gradi Celsius e la sua temperatura media globale alla superficie è di 15 gradi. La serra atmosferica riscalda perciò la superficie di 33 gradi. Marte, invece, irradia a −56 gradi, appena tre gradi meno della sua temperatura superficiale. Il suo effetto serra è attualmente così modesto perché l'atmosfera è tenue.

Affinché quattro miliardi di anni fa la temperatura su Marte superasse il punto di congelamento dell'acqua, dovevano essere disponibili nell'atmosfera maggiori quantità di gas «da serra». Tra questi i candidati più probabili sono anidride carbonica e vapor acqueo; i modelli teorici indicano in particolare che sarebbero stati necessari almeno 1000 millibar di anidride carbonica. Ammesso che una tale quantità di anidride carbonica sia fuoriuscita per degassamento, dove è andata a finire?

Grandi quantitativi di gas potrebbero essere stati assorbiti nel regolite di Marte, ossia lo strato superficiale di suolo a granulosità fine formato dal bombardamento meteoritico. Negli anni settanta Fraser P. Fanale e William A. Cannon del Jet Propulsion Laboratory del California Institute of Technology hanno dimostrato che l'anidride carbonica aderisce a granuli del suolo simili a quelli che presumibilmente costituiscono il regolite. L'anidride carbonica viene anche immagazzinata sotto forma di ghiaccio nelle calotte polari, ma la maggior parte di questo materiale sublima (passa direttamente dallo stato solido a quello di vapore) ogni estate. Sebbene possa esserci una calotta permanente residua al polo sud - i moduli orbitali dei Viking vi hanno fotografato ghiaccio nel 1977, durante l'estate meridionale - la quantità immagazzinata in questo serbatoio sembra piuttosto ridotta in confronto a quella fuoriuscita per degassamento.

Un'altra possibilità è che anche su Marte molta anidride carbonica si trovi nello stesso tipo di serbatoio che contiene la maggior parte dell'anidride carbonica terrestre: le rocce carbonatiche. Per tutto il periodo in cui Marte fu sufficientemente caldo da consentire la presenza di acqua allo stato liquido, questa estrasse anidride carbonica dall'atmosfera e la trasformò in carbonati, così come fanno attualmente gli oceani terrestri. Non è chiaro se vi siano mai state precipitazioni intense su Marte, ma la formazione di carbonati può aver luogo anche senza pioggia, dato che l'anidride carbonica si scioglie bene anche in corpi idrici aperti.

Esiste tuttavia una differenza fondamentale tra Marte e la Terra per quanto concerne la destinazione finale dell'anidride carbonica. Sulla Terra, parte del gas viene costantemente riciclato nell'atmosfera per effetto dei movimenti tettonici delle zolle. Quando due delle zolle in movimento della crosta terrestre collidono, una di esse si immerge nel mantello portando con sé i sedimenti di carbonati. Nel mantello, il calore e la pressione elevati liberano l'anidride carbonica dai sedimenti. Il gas può quindi sfuggire attraverso i vulcani nell'atmosfera.

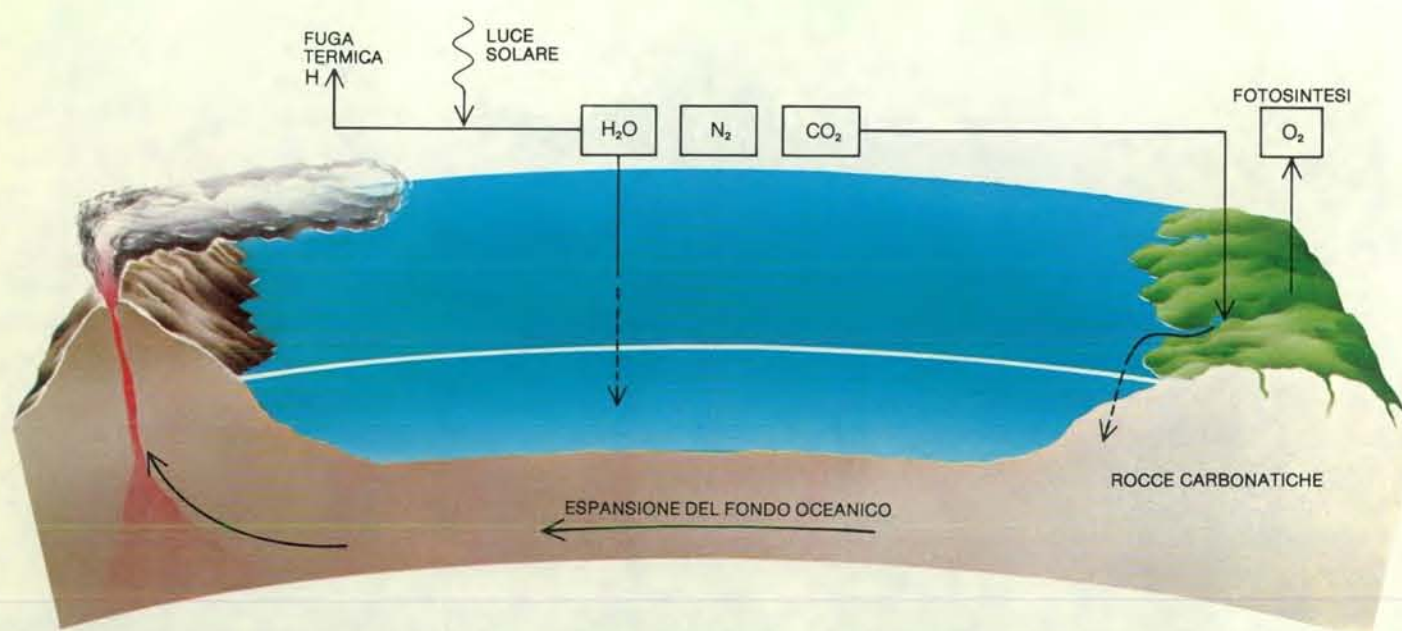
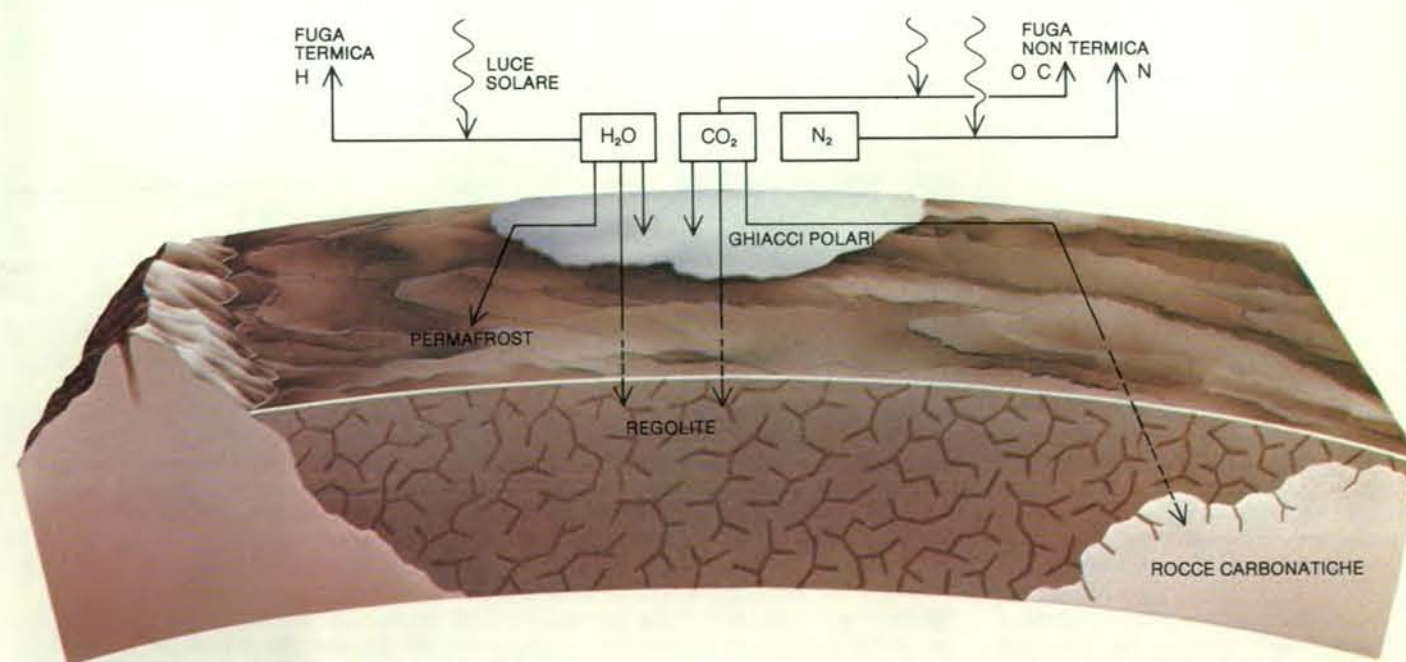
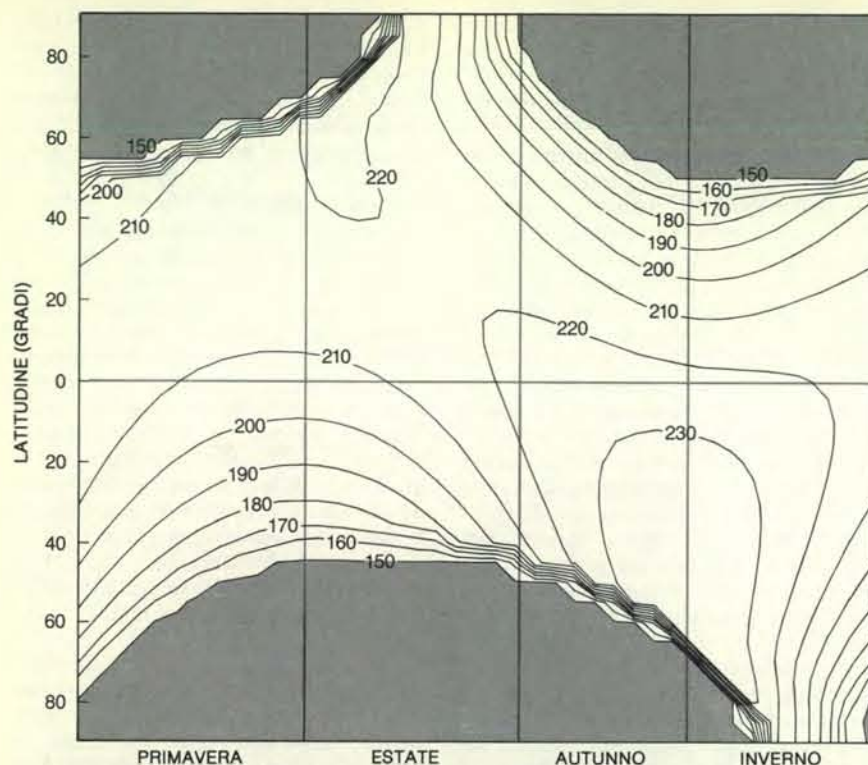
Dal momento che Marte è vulcanicamente meno attivo della Terra, su Marte viene riciclata una quantità inferiore di anidride carbonica. In particolare, non vi è traccia di attività tettonica di più zolle; l'intera crosta appare costituita da un'unica zolla. Il motivo sta nel fatto che Marte è, per dimensioni, solo poco più grande della metà della Terra. La dimensione di un pianeta determina in generale l'intensità delle fonti interne di calore - principalmente processi di decadimento radioattivo ed energia gravitazionale liberata nella fase di accrescimento - che alimentano l'attività tettonica e vulcanica.

Se sulla Terra non vi fosse alcun riciclaggio vulcanico, i processi di alterazione dovuti agli agenti atmosferici trasformerebbero la piccola quantità atmosferica di anidride carbonica in carbonati nel giro di alcuni millenni. Su Marte, un'atmosfera di anidride carbonica corrispondente a una pressione di 1000 millibar potrebbe essersi conservata più a lungo: da 10 a 100 milioni di anni senza riciclaggio, secondo i calcoli di James B. Pollack dello Ames Research Center della NASA. Ma neppure 100 milioni di anni sono sufficienti a spiegare la formazione dei sistemi di valli dovuta all'azione di acqua allo stato liquido. Di conseguenza Pollack ha suggerito che nei primi tempi della sua storia Marte potrebbe avere avuto una quantità di energia residua dovuta all'accrescimento sufficiente ad alimentare un vulcanismo su scala globale. L'anidride carbonica così riciclata, egli ritiene, potrebbe aver dato luogo a un intenso effetto serra per i circa 500 milioni di anni durante i quali furono scavate le valli.

Con la diminuzione dell'attività vulcanica, il livello di anidride carbonica nell'atmosfera e quindi l'effetto serra diminuirono. Infine il pianeta divenne così freddo che l'acqua allo stato liquido non era più stabile. In quella fase la pressione atmosferica in superficie doveva essere probabilmente ben più elevata del livello attualmente osservato. Secondo Ralph A. Kahn della Washington University a St. Louis, su Marte la pressione in superficie è adesso così bassa perché l'anidride carbonica viene continuamente rimossa dall'atmosfera e immagazzinata sotto forma di carbonati in «sacche» transitorie di acqua allo stato liquido. Kahn asserisce che l'esistenza di tali sacche potrebbe essersi prolungata molto, dopo che la temperatura globale media era scesa al di sotto del punto di congelamento; più specificamente, esse potevano formarsi finché la pressione in superficie fosse stata sufficientemente elevata da limitare l'evaporazione. Mediante l'azione di queste sacche transitorie d'acqua la pressione si ridusse gradualmente al valore attuale di 6,1 millibar.

Paradossalmente, dunque, l'acqua allo stato liquido potrebbe in ultima analisi essere stata la causa della propria distruzione. Può darsi che parte di essa sia immagazzinata nei ghiacci polari o in un *permafrost* molto disperso. Il regolite può formare legami chimici o fisici anche con l'acqua, così come si pensa avvenga con l'anidride carbonica.

Se questo scenario è corretto, le ridotte dimensioni di Marte hanno avuto sul clima perlomeno la stessa influenza che ha avuto la sua distanza dal Sole. Inoltre le dimensioni di Marte hanno determinato non solo la sorte dell'anidride carbonica e dell'acqua, ma anche quella dell'azoto, relativamente scarso. Su Marte il livello inferiore di attività vulcanica ha fatto sì che fuoriuscisse per degassamento meno azoto che non sulla Terra. Marte è più piccolo, la sua gravità è più debole, e perciò l'azoto può sfuggire dal pianeta. Sebbene l'azoto non abbia energia termica sufficiente per sfuggire, può acquisire l'energia necessaria mediante il processo di ricombinazione dissociativa. Nell'atmosfera superiore, la radiazione ultravioletta estrae gli elettroni dalle molecole di azoto; quando gli elettroni e l'azoto ionizzato si ricombinano, vi è più energia di quanta sarebbe necessaria per dividere le molecole nei loro atomi costituenti. L'eccedenza fornisce agli atomi energia cinetica sufficiente per sfuggire.

Per quanto riguarda Marte una stima della quantità di azoto liberato per degassamento può essere ottenuta dalle abbondanze relative di azoto 15 e di azoto 14. Poiché l'azoto 15 è più pesante e quindi ha minore probabilità di sfuggire, la sua concentrazione è gradualmente aumentata rispetto a quella dell'azoto 14. I modelli di questa evoluzione partono dal presupposto che Marte e la Terra avessero inizialmente la stessa abbondanza isotopica relativa tuttora rilevata sulla Terra, dalla quale nemmeno l'azoto 14 è in grado di sfuggire. I modelli indicano che è fuoriuscita da Marte una quantità di azoto cor-

Forse le atmosfere primordiali di Marte e della Terra erano simili, ma poi hanno subito una diversa evoluzione. Su entrambi i pianeti i componenti predominanti dell'atmosfera primordiale erano anidride carbonica, vapore acqueo e azoto liberati per degassamento vulcanico o da impatto. Sulla Terra (*qui sopra*) l'azoto è rimasto nell'atmosfera e la maggior parte del vapore acqueo si è condensata negli oceani. L'anidride carbonica è stata invece rimossa dall'atmosfera dalle precipitazioni e da processi di alterazione e poi immagazzinata nelle rocce carbonatiche. Una certa quantità di anidride carbonica, tuttavia, ritorna all'atmosfera attraverso i vulcani. Su Marte (*a destra*) il tasso di vulcanismo è ora così basso da rendere quasi impossibile il riciclaggio di anidride carbonica. Di conseguenza, i processi di alterazione dovuti all'acqua allo stato liquido hanno distrutto l'atmosfera di anidride carbonica che per effetto serra manteneva la temperatura del pianeta abbastanza elevata da consentire la presenza di acqua allo stato liquido. La maggior parte dell'anidride carbonica è probabilmente immagazzinata nei carbonati oppure adsorbita nel suolo a granulometria fine del regolite. Il polo nord ha una calotta di ghiaccio d'acqua permanente, ma la maggior parte dell'acqua deve trovarsi nel regolite o in un permafrost sparso. L'azoto è sfuggito alla debole gravità marziana traendo energia da reazioni fotochimiche.

Le temperature superficiali di Marte seguono un ciclo stagionale, ma le stagioni sono ineguali per durata e intensità a causa della marcata eccentricità dell'orbita marziana. (Le stagioni cui si riferiscono questa illustrazione e le successive sono quelle dell'emisfero settentrionale.) Le isolinee sono isoterme, ossia linee di uguale temperatura espressa in kelvin (gradi Celsius sopra lo zero assoluto, dove lo 0 gradi Celsius corrisponde a 273 kelvin). L'anidride carbonica solidifica a circa 150 kelvin, e quindi le isolinee di 150 gradi indicano approssimativamente i margini delle calotte polari (*in grigio*). La calotta meridionale è più estesa della settentrionale perché al sud l'inverno e l'autunno sono più lunghi e più freddi. Le temperature massime si hanno nella tarda primavera (autunno settentrionale) nella fascia subtropicale meridionale. Quando le calotte glaciali si ritirano, il gradiente termico latitudinale lungo i margini diventa ripido. La cartina è stata realizzata sulla base di un modello sviluppato da Hugh H. Kieffer dell'United States Geological Survey.

rispondente a una pressione di anche 30 millibar, quasi 200 volte la quantità attualmente presente nell'atmosfera. La concentrazione attuale è a sua volta solo una minuscola frazione della concentrazione che si ha nell'atmosfera terrestre.

Sebbene la Terra e Marte abbiano sviluppato atmosfere di composizione molto diversa, i due pianeti hanno alcune caratteristiche climatiche in comune. Per esempio, l'andamento della circolazione generale è analogo e questo per diverse ragioni.

In primo luogo, essendo pressoché trasparenti alla luce del Sole, l'atmosfera marziana e quella terrestre vengono entrambe riscaldate dal basso. In secondo luogo, su entrambi i pianeti i tropici ricevono una quantità di energia solare considerevolmente maggiore rispetto ai poli; il gradiente termico risultante produce un gradiente di pressione che alimenta una circolazione meridionale (nel senso nord-sud). In terzo luogo, i pianeti ruotano rapidamente e con velocità pressoché uguali. (Un giorno su Marte è circa 40 minuti più lungo di un giorno terrestre.) La velocità di rotazione determina l'entità della forza di Coriolis, una forza apparente che fa deviare una massa d'aria in moto.

I modelli teorici realizzati da Conway B. Leovy dell'Università di Washington, da Pollack e da me hanno messo in luce le caratteristiche principali della circolazione generale marziana. Come avviene anche sulla Terra, la circolazione meridionale prende la forma di due celle di Hadley, una in ciascun emisfero. Le celle condividono un ramo ascendente in prossimità dell'equatore, dove il riscaldamento solare è massimo. L'aria sale in questa zona, si muove verso i poli negli strati superiori, quindi si raffredda e scende nelle regioni subtropicali (a latitudini di circa 25 gradi). Infine ritorna verso i tropici in prossimità della superficie. Poiché la forza di Coriolis devia l'aria in movimento verso destra nell'emisfero settentrionale e verso sinistra nell'emisfero meridionale, la deviazione dell'aria diretta verso i poli nelle celle di Hadley dà luogo a venti occidentali in quota; alla superficie la deviazione dell'aria diretta verso l'equatore produce venti alisei orientali analoghi a quelli terrestri.

Un regime differente di circolazione, pure analogo al suo corrispettivo terrestre, prevale a medie ed elevate latitudini. Qui predominano venti da ovest sia alla superficie, sia in quota. Una stretta fascia di venti intensi, le cosiddette correnti a getto, spira nell'atmosfera superiore, a una quota compresa forse tra 30 e 40 chilometri. In prossimità della superficie, i sistemi di perturbazione costituiti da centri di alta e di bassa pressione migrano verso est, seguendo il percorso della corrente a getto sovrastante e traendo da essa energia. I venti ciclonici e anticiclonici che vorticano attorno ai centri di pressione svolgono la funzione espletata dalle celle di Hadley alle latitudini inferiori, ossia trasportano energia verso i poli.

A differenza delle perturbazioni terrestri, i cui movimenti sono notoriamente difficili da prevedere, i sistemi meteorologici marziani sono di una regolarità perfino monotona. Nei punti di atterraggio dei moduli Viking le perturbazioni si susseguivano di norma a intervalli compresi tra due e quattro giorni. La causa di tale regolarità è incerta, ma può essere messa in relazione con il fatto che l'anidride carbonica, il componente principale dell'atmosfera marziana, assorbe ed emette rapidamente radiazione termica. Può darsi che un'atmosfera con estesa capacità radiativa ammortizzi rapidamente le variazioni termiche provocate dalle perturbazioni. Le perturbazioni terrestri possono invece essere più irregolari perché l'azoto, il componente principale dell'atmosfera terrestre, è radiativamente inerte.

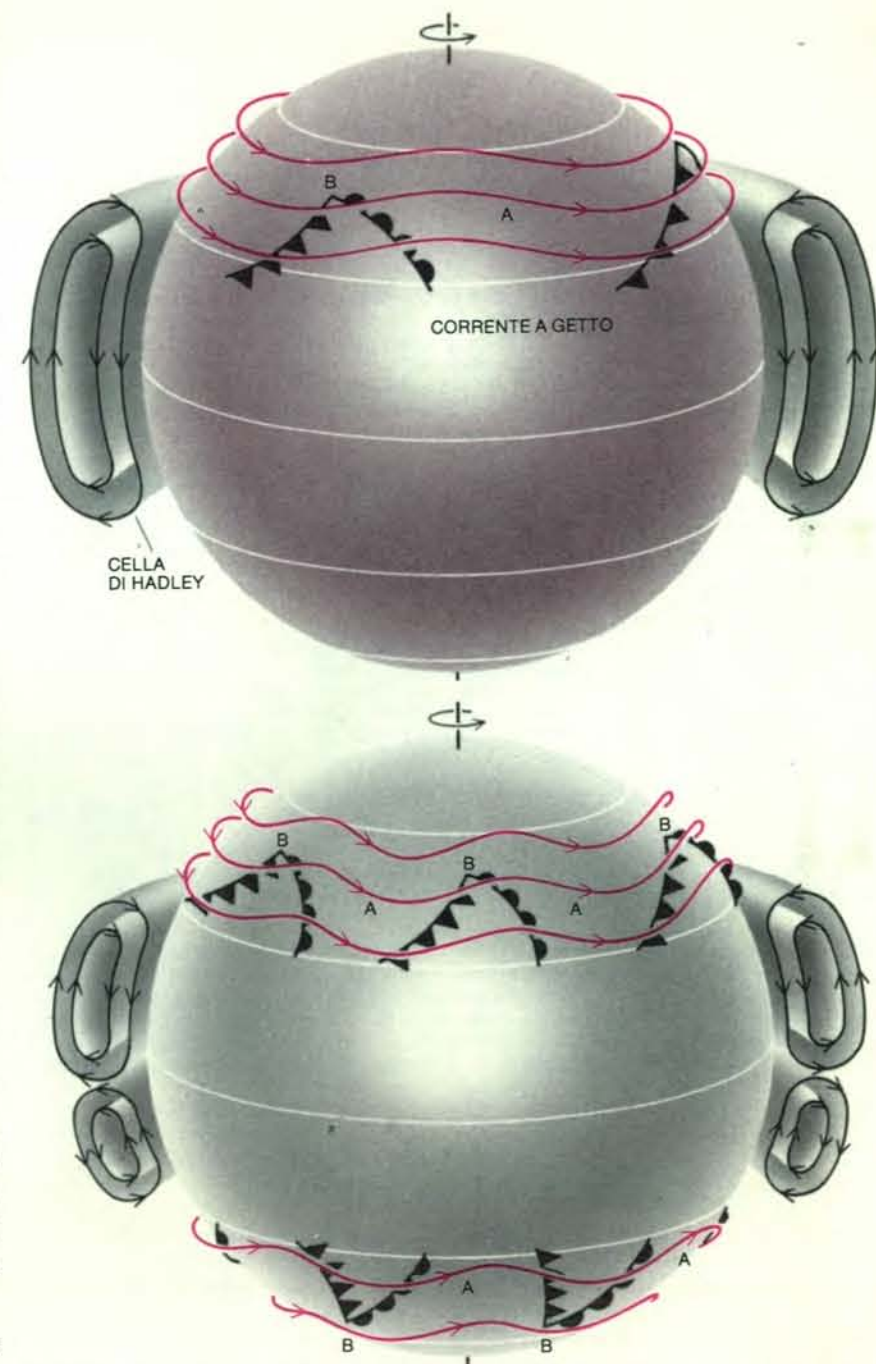
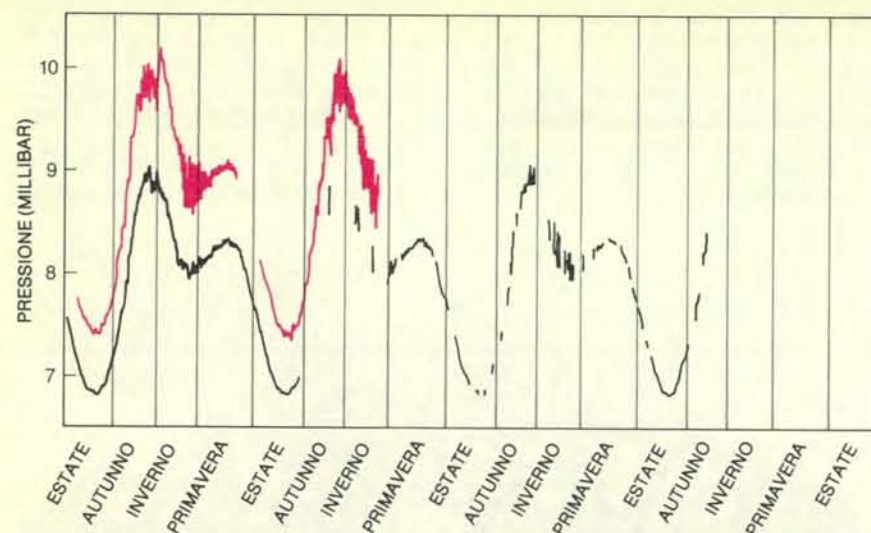
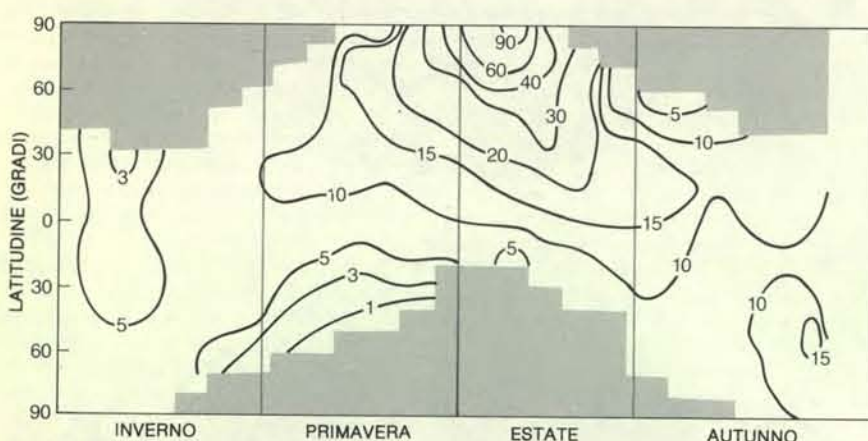
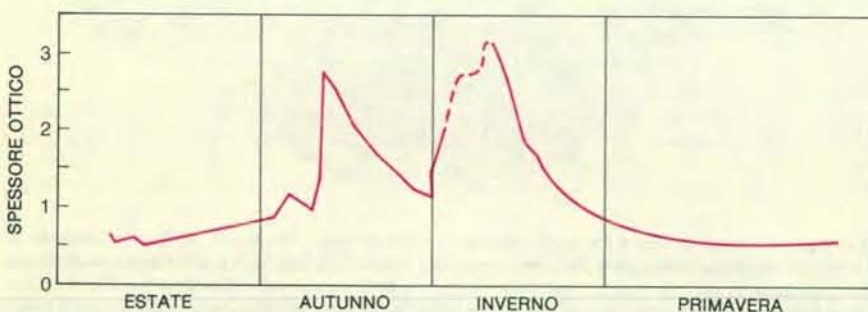
La circolazione marziana cambia stagionalmente, come avviene sulla Terra, e per la stessa ragione. Un pianeta ha stagioni se il suo asse di rotazione non è perpendicolare al piano orbitale; quando il pianeta si muove attorno al Sole, l'inclinazione del suo asse porta a una variazione nella distribuzione latitudinale della radiazione solare incidente e a una variazione corrispondente nella distribuzione delle temperature: maggiore è l'angolo di inclinazione (l'obliquità), più pronunciati sono i cambiamenti stagionali. L'angolo è attualmente quasi uguale per la Terra (23,5 gradi) e per Marte (25,2 gradi), e pertanto le variazioni stagionali sui due pianeti sono simili. Vi sono tuttavia due fattori che determinano su Marte un andamento stagionale leggermente diverso: Marte non ha oceani, e la sua orbita è più ellittica di quella terrestre.

Un modo in cui su Marte si manifesta l'assenza di oceani è attraverso variazioni stagionali nella configurazione delle celle di Hadley. Il ramo ascendente delle celle di Hadley tende a essere sopra il punto più caldo del pianeta. Poiché il suolo marziano, a differenza degli oceani terrestri, risponde rapidamente alle variazioni quantitative della radiazione solare incidente (la cosiddetta insolazione), il punto più caldo segue di norma fedelmente il percorso annuale del Sole. Ai solstizi il ramo ascendente può pertanto trovarsi anche a 25 gradi di latitudine dall'equatore. I modelli teorici suggeriscono che le due celle di Hadley siano allora sostituite da una singola cella transequatoriale. Questo non avviene mai sulla Terra; il ramo ascendente in comune alle due celle di Hadley segue solo parzialmente il cammino del Sole nelle sue escursioni primaverili e autunnali rispetto all'equatore, poiché gli oceani rispondono solo lentamente alle variazioni di temperatura.

L'assenza di oceani influisce anche sulla circolazione alle medie latitudini. L'intensità di una corrente a getto dipende dal grado in cui varia con la latitudine la temperatura dell'atmosfera sottostante: maggiore la variazione, maggiore il gradiente di pressione in quota e più forte il getto. Sulla Terra gli oceani proteggono il gradiente termico latitudinale da forti variazioni stagionali; le correnti a getto spirano dunque durante l'intero arco dell'anno, sebbene la loro intensità sia massima in inverno. La situazione su Marte è diversa. In inverno e in primavera la corrente a getto marziana è quasi quattro volte più intensa delle correnti a getto terrestri, poiché in queste stagioni la calotta glaciale polare si estende fino alle medie latitudini. Il contrasto termico tra il ghiaccio di anidride carbonica e il suolo circostante è particolarmente forte in primavera, allorché il suolo in prossimità del ghiaccio si riscalda rapidamente. In estate, tuttavia, la corrente a getto scompare, perché in quel periodo la calotta glaciale si è ritirata e il suolo ha più o meno la stessa temperatura in tutto l'emisfero. Lo stesso andamento è seguito dai sistemi di perturbazione che traggono la loro energia dalla corrente a getto.

La seconda differenza fondamentale tra la Terra e Marte per quanto riguarda le stagioni è che l'orbita di Marte è più eccentrica, ovvero più ellittica, dell'orbita della Terra. Di conseguenza le stagioni marziane sono di durata e di intensità ineguali. (Sono più lunghe delle stagioni terrestri dato che l'anno marziano ha una durata di 687 giorni terrestri.) Il pianeta raggiunge la minima distanza dal Sole (perielio) nella tarda primavera meridionale, che è 52 giorni più corta della primavera settentrionale, quando il pianeta è alla massima distanza dal Sole (afelio). Al perielio, Marte riceve circa il 40 per cento di radiazione solare in più che non all'afelio; la corrispondente differenza sulla Terra è solo del 3 per cento. Rispetto alle stagioni settentrionali, primavera e estate meridionali sono brevi e calde, mentre autunno e inverno sono freddi e lunghi. Questa asimmetria ha un notevole influsso su tre cicli stagionali interagenti: il ciclo dell'anidride carbonica, il ciclo dell'acqua e quello della polvere.

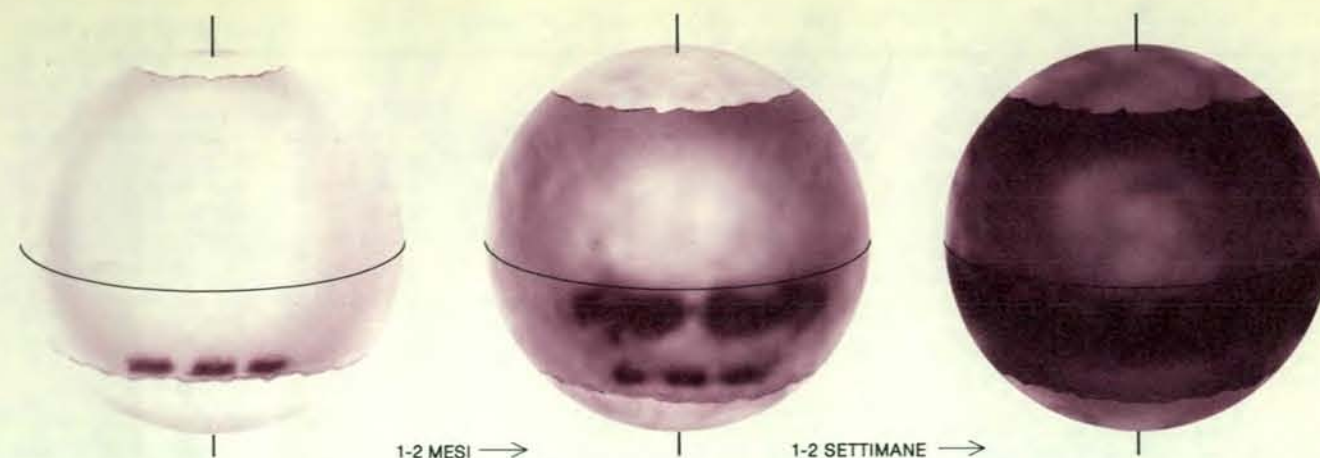
Durante l'inverno, alle alte latitudini, la temperatura su Marte scende al punto da far solidificare l'anidride carbonica, un fenomeno questo esclusivamente marziano. Ogni anno, su Marte, circa il 20 per cento dell'atmosfera entra ed esce ciclicamente nelle calotte polari, provocando una variazione su scala planetaria della pressione in superficie. In effetti, l'atmosfera subisce oscillazioni su base stagionale, accumulandosi sul polo invernale. Dal momento che nell'emisfero meridionale autunno e inverno durano più a lungo, la calotta polare meridionale si estende maggiormente della calotta polare settentrionale. Nella sua massima estensione raggiunge una latitudine di circa 45 gradi, mentre la calotta settentrionale raggiunge una latitudine di 50 gradi.

Il ciclo dell'anidride carbonica non solo determina le variazioni stagionali della pressione in superficie, ma influisce anche sul comportamento stagionale dell'acqua nell'atmosfera marziana. Allorché l'estate subentra alla primavera le calotte polari si ritraggono in ciascun emisfero, ma non scompaiono mai completamente. La com-

Il sistema dei venti su Marte (*in alto*) è simile a quello terrestre (*in basso*). Su Marte, tuttavia, le variazioni stagionali del regime dei venti sono più pronunciate perché il suolo marziano risponde più rapidamente degli oceani terrestri alle variazioni dell'energia solare incidente. L'illustrazione mostra la circolazione generale dell'atmosfera durante l'inverno settentrionale. Sulla Terra, il ramo ascendente delle celle di Hadley tropicali è situato leggermente a sud dell'equatore. Su Marte il ramo ascendente si sposta di circa 25 gradi a sud dell'equatore; le due celle di Hadley, distinte durante l'autunno, formano poi una singola cella transequatoriale. Su entrambi i pianeti la circolazione alle medie latitudini è dominata da correnti a getto occidentali di alta quota che sospingono le perturbazioni superficiali. Su Marte, però, il gradiente termico latitudinale che alimenta le correnti a getto è forte solo nell'emisfero invernale; in estate getti e perturbazioni scompaiono.

Su Marte il ciclo dell'anidride carbonica è stato documentato dai moduli di atterraggio dei Viking, i quali hanno registrato fluttuazioni stagionali della pressione atmosferica in superficie. La pressione tende a salire quando l'anidride carbonica sublima dai poli e a diminuire quando l'anidride carbonica solidifica, ma poiché la calotta polare meridionale si estende di più di quella settentrionale, le sue fluttuazioni predominano. *Viking 2* (*in colore*) ha rilevato valori di pressione maggiori di *Viking 1* perché era a una quota più bassa. I dati sono stati elaborati da James E. Tillman.

Il ciclo dell'acqua su Marte dipende dal ciclo dell'anidride carbonica. I moduli orbitali dei Viking hanno misurato variazioni stagionali nella distribuzione latitudinale del vapore acqueo; la sua abbondanza nell'atmosfera è data in termini di profondità (micrometri) che l'acqua avrebbe se fosse allo stato liquido. L'abbondanza massima è stata osservata al polo nord, durante l'estate, quando il ghiaccio di anidride carbonica era completamente sublimato lasciando scoperto il ghiaccio d'acqua. Delle aree in grigio non si hanno rilievi. I dati sono stati elaborati da Bruce M. Jakosky.

Il tracciato del ciclo della polvere è stato effettuato dal modulo di atterraggio di *Viking 1* sulla base del grado di attenuazione della luce solare attraverso l'atmosfera. Quanto maggiore è la quantità di polvere presente, tanto maggiore è l'opacità, o spessore ottico, dell'atmosfera. L'atmosfera marziana non è mai perfettamente tersa. I massimi di opacità in autunno e in inverno indicano la comparsa delle tempeste globali di polvere. Il tratteggio si riferisce a un periodo in cui si è rilevato solo un limite inferiore di opacità. I dati sono stati elaborati da James B. Pollack e collaboratori.

posizione delle calotte residue era incerta finché i moduli orbitali dei Viking non le sorvolarono nel 1976 e nel 1977. Si è rilevato che la calotta residua settentrionale è di ghiaccio d'acqua, mentre la calotta residua meridionale è composta prevalentemente di anidride carbonica. Dalle misurazioni della luce solare riflessa i moduli orbitali dei Viking hanno potuto tracciare la distribuzione areale del vapore d'acqua nell'atmosfera. L'abbondanza massima di vapore acqueo si è riscontrata durante l'estate nella regione polare settentrionale; nessun picco analogo è stato osservato durante l'estate al polo sud.

Il fenomeno ha una spiegazione semplice. La calotta permanente di ghiaccio d'acqua al polo nord diventa una sorgente di vapore acqueo atmosferico ogni estate, allorché la sublimazione del ghiaccio di anidride carbonica sovrastante la espone. Se, invece, al polo sud il ghiaccio di anidride carbonica sopravvive sempre all'estate, il vapore acqueo a contatto con esso gela e viene rimosso dall'atmosfera. In altre parole, la calotta polare meridionale sarebbe un pozzo di vapore acqueo che cresce a spese della calotta settentrionale. Sul lungo periodo quest'ultima non sarebbe affatto permanente, ma scomparirebbe.

Non è tuttavia certo che al polo sud il ghiaccio di anidride carbonica sopravviva indefinitamente estate dopo estate. Osservazioni telescopiche di Marte compiute nel 1969, quando il pianeta era vicino alla Terra, avevano rilevato, durante l'estate meridionale, una quantità di vapore acqueo atmosferico molto maggiore di quella rilevata dai Viking durante la stessa stagione nel 1977. Può darsi che al polo sud, nel 1969, la calotta di anidride carbonica sia sublimata esponendo una calotta di ghiaccio d'acqua, analoga a quella del polo nord. Che la calotta di anidride carbonica meridionale sia permanente o meno, il fatto che riesca a sopravvivere almeno ad alcune estati, a differenza della calotta settentrionale, richiede una spiegazione, che probabilmente ha a che fare con il terzo ciclo stagionale marziano: il ciclo della polvere.

Misurando l'attenuazione della luce solare nell'atmosfera marziana, i moduli di atterraggio dei Viking hanno potuto rilevare il comportamento della polvere atmosferica nell'arco di un ciclo completo di stagioni. L'atmosfera non è mai perfettamente tersa; la limpidezza massima si ha durante la primavera e l'estate settentrionali, ma anche allora il cielo fotografato dai moduli di atterraggio aveva il colore giallo-rosato tipico della luce solare diffusa da particelle in sospensione. Nell'autunno e nell'inverno settentrionali il brusco aumento dell'opacità indica l'arrivo delle tempeste di polvere di scala planetaria. Tali tempeste sono così intense che le particelle di polvere vengono sollevate a oltre 40 chilometri di altezza, mentre la morfologia della superficie risulta del tutto oscurata. Dato che su Marte le precipitazioni in grado di ripulire l'atmosfera sono scarse o nulle, la polvere può rimanere in sospensione per settimane o anche per mesi.

Le tempeste globali di polvere avvengono nell'inverno e nell'autunno settentrionali quando il riscaldamento di Marte da parte del Sole è più intenso. Mentre durante la primavera meridionale Marte si avvicina al perielio (*a sinistra*), lungo il margine della calotta polare si sviluppano tempeste locali. La polvere così sollevata si diffonde in tutta l'atmosfera che, assorbendo più radiazione solare, si riscalda, con conseguente intensificazione dei venti di marea e della circolazione di Hadley. L'effetto è particolarmente intenso nella fascia subtropicale meridionale, dove si sviluppano tempeste locali di maggiore portata (*al centro*), che sollevano più polvere e riscaldano ulteriormente l'atmosfera. Alcuni giorni dopo, i venti di marea diventano abbastanza intensi da sollevare polvere in tutta la fascia subtropicale meridionale. La polvere pervade il ramo ascendente della cella di Hadley transequatoriale diffondendosi nell'emisfero settentrionale (*a destra*). La tempesta si placa quando l'atmosfera è talmente carica di polvere che i suoi strati inferiori non vengono più riscaldati dalla luce solare, i venti superficiali cadono e cessa il sollevamento di polvere.

Le tempeste globali di polvere avvengono in modo particolare durante l'autunno e l'inverno settentrionali (primavera ed estate meridionali), quando Marte è prossimo al perielio, perché è questo il momento in cui il riscaldamento della sua atmosfera e il vigore della sua circolazione sono massimi. Secondo una teoria proposta nel 1973 da Leovy, Richard W. Zurek (ora al Jet Propulsion Laboratory) e Pollack, le tempeste globali sono attribuibili a un effetto di retroazione positiva nell'atmosfera tra i venti di superficie e la polvere. Questa, sollevata dai venti, tende a intensificarli poiché assorbe radiazione solare e quindi riscalda direttamente l'atmosfera.

Secondo questa teoria, i venti alimentati dal forte gradiente termico lungo il bordo della calotta polare meridionale che si sta ritirando generano numerose tempeste di piccola entità che gradualmente fanno aumentare il livello generale di polverosità atmosferica. La polvere a sua volta intensifica la circolazione di Hadley come pure i cosiddetti venti di marea. (Questi venti vengono generati dal ciclo quotidiano di riscaldamento: l'aria sale sul lato diurno caldo del pianeta e rifluisce verso il lato notturno.) La forza combinata dei venti è massima nella fascia subtropicale meridionale, dove è massima l'insolazione. La topografia di questa zona favorisce la formazione di perturbazioni addizionali che sollevano altra polvere nell'atmosfera.

Una volta raggiunto un livello critico di polverosità, la retroazione positiva subisce una rapida accelerazione. La circolazione di Hadley si espande e diffonde polvere verso l'equatore e oltre. Una simulazione al calcolatore da me realizzata ha mostrato come dopo sei giorni di sollevamento di polvere, la nube può diffondersi nell'emisfero settentrionale. La retroazione continua fintantoché l'atmosfera non è così carica di polvere da far sì che i suoi strati inferiori ricevano poco calore e di conseguenza i venti di superficie cadano.

Il ciclo della polvere può ripercuotersi sui cicli dell'anidride carbonica e dell'acqua in un modo che dovrebbe aiutare a spiegare le differenze tra la calotta polare meridionale e quella settentrionale. Dal momento che le tempeste globali di polvere avvengono nell'inverno settentrionale ma non nell'inverno meridionale, è naturale supporre che la calotta settentrionale di anidride carbonica incorpori più polvere durante la sua formazione. Poiché una calotta di ghiaccio sporco assorbe più radiazione solare e ne riflette meno di una pulita, tutta l'anidride carbonica potrebbe sublimare dalla calotta settentrionale fin dall'inizio dell'estate, mentre la calotta meridionale, più pulita e più brillante, sopravviverebbe. Un'analisi dei dati dei Viking effettuata da David Paige del Jet Propulsion Laboratory dimostra che la calotta meridionale è effettivamente più brillante di quella settentrionale.

Purtroppo altri dati dei Viking inficiano questa ipotesi così immediata. Su Marte le tempeste globali di polvere non avvengono ogni anno; esse sono state osservate in soli due dei quattro anni di attività del modulo di atterraggio di *Viking 1*. Se l'ipotesi fosse corretta, la variazione stagionale nel livello atmosferico di anidride carbonica, e quindi nella pressione in superficie, dovrebbe essere diversa in un anno di tempeste globali rispetto alla variazione registrata in un anno senza tempeste. In un anno senza tempeste la calotta settentrionale sarebbe presumibilmente pulita, e potrebbe anche sopravvivere all'estate come fa la calotta meridionale. La pressione in superficie durante la primavera settentrionale risulterebbe perciò più bassa di quanto sarebbe se l'intero contenuto di anidride carbonica della calotta fosse sublimato nella atmosfera. Tuttavia i dati dei Viking non mostrano alcuna differenza significativa nella variazione stagionale di pressione da un anno all'altro. Evidentemente il ciclo dell'anidride carbonica non è direttamente governato dalla presenza o dall'assenza di tempeste globali di polvere.

Paige ha suggerito che la polvere potrebbe nondimeno spiegare la scomparsa annuale della calotta settentrionale e la sopravvivenza di quella meridionale. Mediante una accurata analisi dei dati dei Viking egli ha scoperto che le calotte diventano più brillanti con l'avvicinarsi dell'estate, quando l'intensità della luce solare che giunge alla loro superficie aumenta. Il motivo, egli suggerisce, è che le particelle di polvere sulla superficie del ghiaccio divengono sufficientemente calde da provocare la sublimazione dei granuli di anidride carbonica che le trattengono. Le particelle di polvere affondano nel ghiaccio, lasciando la superficie più pulita e brillante.

La riflettività della calotta dipende perciò non tanto dalla quantità di polvere depositata, quanto dalla profondità raggiunta dalle particelle, e questa a sua volta dalla quantità di energia solare che incide sulla calotta in primavera e in estate. Poiché Marte è più vicino al Sole durante l'estate meridionale che durante quella settentrionale, la calotta meridionale è sempre più brillante (più riflettente) di quella settentrionale indipendentemente dalle variazioni annuali di polverosità dell'atmosfera. In altre parole, la calotta meridionale sopravvive all'estate proprio perché l'estate meridionale è intensa. Una calotta che si autoprotegge dal Sole dovrebbe avere un significativo influsso sulla stabilità a lungo termine del clima di Marte.

La presenza nell'emisfero meridionale di Marte di un sistema di valli dimostra che il clima era un tempo abbastanza caldo perché vi fosse su Marte acqua allo stato liquido. La ramificazione dei canali, la cui età è stimata intorno ai quattro miliardi di anni, suggerisce che essi siano stati formati da corsi convergenti; le dimensioni ridotte indicano che la loro portata era scarsa. L'area ripresa nell'immagine è larga 250 chilometri.

Vi sono buoni motivi per ritenere che il clima marziano muti ciclicamente, poiché al pari della Terra Marte è soggetto a variazioni periodiche di tre parametri orbitali: l'eccentricità, l'obliquità (inclinazione dell'asse di rotazione) e la direzione dell'asse, che descrive un lento moto di precessione. Ciascuna di queste variazioni può influenzare il clima cambiando la distribuzione stagionale e latitudinale dell'energia solare incidente. Un incremento di eccentricità fa aumentare la differenza dell'intensità dei fenomeni stagionali tra i due emisferi; un incremento dell'obliquità espone i poli a maggiori quantità di luce solare; la precessione dell'asse determina la stagione nella quale il pianeta raggiunge il perielio.

È largamente diffusa la convinzione che le variazioni orbitali della Terra costituiscano il motivo delle oscillazioni del suo clima nel milione di anni trascorso tra epoche glaciali e periodi caldi. Le variazioni orbitali di Marte sono molto maggiori di quelle terrestri. (Esse hanno anche periodi più lunghi, di un ordine compreso tra 100 000 e un milione di anni.) L'obliquità dell'asse di Marte, per esempio, che attualmente è di 25,2 gradi, può variare di circa 13 gradi in entrambi i sensi rispetto al valore attuale. Invece, l'obliquità dell'asse della Terra varia solo di circa un grado.

Il principale effetto climatico delle variazioni dell'orbita marziana sarebbe costituito dalla variazione della pressione atmosferica in superficie in seguito al variare della quantità di anidride carbonica nell'atmosfera. È probabile, per esempio, che si formi una calotta pressoché perenne di ghiaccio di anidride carbonica su almeno uno dei poli durante periodi di scarsa obliquità, quando i poli ricevono poca luce solare. La pressione in superficie durante tali periodi sarebbe relativamente bassa. Le tempeste di polvere sarebbero rare o inesistenti, dal momento che quando la pressione in superficie è bassa sono necessari venti molto più forti per sollevare polvere dalla superficie. In periodi di elevata obliquità, invece, la calotta di ghiaccio di anidride carbonica sublimerebbe e l'anidride carbonica potrebbe anche essere trasferita dal regolite nell'atmosfera. La pressione in superficie aumenterebbe e le tempeste di polvere diventerebbero frequenti.

Mariner 9 e i moduli orbitali dei Viking potrebbero avere scoperto prove di tali variazioni. A entrambi i poli i moduli orbitali hanno fotografato terreni stratificati la cui età è stata stimata di almeno alcuni milioni di anni. L'alternanza e lo spessore pressoché uniformi degli strati suggeriscono che essi si siano formati per processi atmosferici di sedimentazione (deposizione di polvere atmosferica) regolati da periodiche variazioni climatiche. Owen B. Toon dello Ames Research Center e collaboratori ritengono che gli strati siano falde sepolte di ghiaccio d'acqua. Secondo la loro ipotesi, ogni strato si è depositato durante un periodo di scarsa obliquità, nel quale le calotte di ghiaccio di anidride carbonica fecero gelare il vapore acqueo che era stato trasferito dal permafrost nell'atmosfera. Durante il successivo periodo di elevata obliquità, la polvere sollevata da frequenti tempeste globali seppellì la falda prima che tutto il ghiaccio d'acqua potesse evaporare. Perciò, l'impilamento degli strati visibile oggi potrebbe essere stato prodotto da una successione di cicli di obliquità.

Nonostante appaia evidente che Marte abbia subito mutamenti climatici ciclici, non è affatto chiaro quanto siano stati pronunciati tali mutamenti. Si è ipotizzato che la pressione atmosferica su Marte potrebbe variare in un intervallo compreso tra meno di un millibar e parecchie volte l'attuale valore di 6,1 millibar. L'entità delle fluttuazioni di pressione, tuttavia, dipende dalla quantità di anidride carbonica immagazzinabile in una calotta di ghiaccio perenne e dalla quantità scambiata invece tra l'atmosfera e il regolite. Nessuna di queste grandezze è nota. Inoltre, se l'ipotesi di Paige è corretta e le calotte di ghiaccio di anidride carbonica marziane proteggono se stesse dagli incrementi di intensità della radiazione solare, allora le oscillazioni di pressione provocate dalla sublimazione del ghiaccio risulterebbero smorzate. Quando *Mars Observer* entrerà nell'orbita polare di Marte, sarà in grado di eliminare alcuni di questi dubbi. Nel corso della sperimentazione esso dovrebbe aiutare i ricercatori a determinare la natura e l'intensità delle variazioni climatiche di Marte.

Rischio genetico per il ghepardo

L'animale terrestre più veloce del mondo è in lotta per la sopravvivenza: una drastica riduzione numerica della sua popolazione ha prodotto uniformità genetica, rendendolo vulnerabile a ogni cambiamento ecologico

di Stephen J. O'Brien, David E. Wildt e Mitchell Bush

Il ghepardo, una vera e propria macchina da corsa, è un modello di ingegneria aerodinamica. Ha un cranio piccolo e leggero e zampe lunghe e snelle, non molto diverse da quelle di un levriere. Cuore, sistema vascolare, polmoni e ghiandole surrenali sono ingrossati e accentuano le capacità dell'animale ad accelerare e a seguire la rotta durante una caccia condotta a elevata velocità. Il ghepardo ha, inoltre, gli artigli semiretrattili: sempre estesi come cunei, a differenza degli artigli delle altre specie di felidi, che sono normalmente racchiusi in un astuccio protettivo. Grazie a questi adattamenti, il ghepardo si comporta da abile cacciatore nelle savane aperte e pianeggianti dell'Africa centrale e meridionale, dove uccide una percentuale di prede più elevata persino di quella del leone. Dopo essersi avvicinato di soppiatto alla vittima prescelta, esso si lancia in una caccia sfrenata, a una velocità cronometrata attorno ai 110 chilometri orari, balza sulla preda o la rovescia a terra e rapidamente la uccide per strangolamento con le forti mascelle da felino.

Nonostante la grande abilità nella corsa e nella caccia, la specie attualmente vivente (*Acinonyx jubatus*) sembra dirigersi verso l'estinzione. Mentre un tempo scorrazzava per tutto il mondo, ora la sua area di diffusione è limitata a poche sacche in Africa. In tutto il mondo sopravvivono forse non più di 20 000 esemplari.

Come si può spiegare questo declino? Dopo cinque anni di ricerche, siamo arrivati alla conclusione che la specie abbia perso per qualche motivo la sua variabilità genetica. A causa di accoppiamenti consanguinei (inincroci) avvenuti alcune generazioni fa si è sviluppata una popolazione di ghepardi quasi identici tra loro. Fin da quando, più di un secolo fa, Charles Darwin scrisse *L'origine delle specie* è stato ampiamente dimostrato che l'uniformità genetica costituisce un notevole ostacolo alla capacità di una specie di adattarsi a perturbazioni ecologiche come le variazioni di temperatura, la siccità, le glaciazioni, l'influenza di nuovi virus o batteri. Secondo la legge darwiniana della selezione naturale, gli individui ben adattati a un ambiente avranno un maggior numero di discendenti di quelli meno adattati. Quando una specie presenta scarsa variabilità genetica, tra i suoi ranghi vi saranno minori probabilità di trovare molti membri con caratteri genetici ben adatti a sopportare modificazioni dell'ambiente; in condizioni mutate la specie ha perciò scarse possibilità di competere in modo efficace per la sopravvivenza e potrebbe estinguersi.

Come lo stesso Darwin avrebbe potuto prevedere, per effetto di accoppiamenti consanguinei il ghepardo ha assunto caratteri di disadattamento talmente gravi da annullare l'ottimo adattamento legato alla snellezza del corpo. La specie è infatti assai vulnerabile alle malattie e presenta una mortalità nelle prime fasi di vita che in alcune riserve è valutata attorno al 70 per cento. Va anche detto che il ghepardo, pur essendo il mammifero più veloce della Terra, può correre solo per alcune centinaia di metri. Poiché durante un tipico inseguimento di caccia il ghepardo aumenta la sua frequenza respiratoria da 60 a 150 atti al minuto, spesso, poi crolla per circa mezz'ora prima di riacquistare le forze. Durante questo intervallo potrebbe essere attaccato o, al minimo, potrebbe perdere la preda. A dire il vero, anche quando i ghepardi non sono esausti, difendono debolmente le prede uccise, lasciandosele strappare per una buona metà da animali più aggressivi, come leoni, leopardi e iene.

Abbiamo cominciato a compiere ricerche sulle cause del declino del ghepardo nel 1981, dopo che Frank Brand, direttore dei National Zoological Gardens of South Africa, ci aveva invitato, quali rappresentanti dello US National Zoological Park, a cooperare nello studio di un problema apparentemente limitato: la difficoltà di allevare il ghepardo in cattività. Tra i primi frutti di questa collaborazione abbiamo potuto rilevare che la specie aveva probabilmente un pool genico impoverito: perciò ci siamo sentiti stimolati a compiere ricerche particolareggiate sulla sua condizione genetica.

Il problema di Brand era sufficientemente chiaro. Nel 1971 gli Zoological Gardens avevano istituito un programma di allevamento del ghepardo al De Wildt Cheetah Breeding and Research Center, un istituto nei pressi di Pretoria. Dieci anni dopo il centro era all'avanguardia nel mondo per il numero di ghepardi che era riuscito a far nascere, ma non era ancora in grado di risolvere i problemi legati alla bassa fecondità e all'alta mortalità dei soggetti in tenera età (il 37 per cento).

In realtà, il ghepardo ha alle spalle tutta una serie di insuccessi nei tentativi condotti per farlo riprodurre in cattività e questa storia risale almeno al tempo di Akbar il Grande, sovrano dell'India nel XVI secolo. (Akbar, che possedeva 1000 ghepardi, era uno dei tanti sovrani del passato che avevano mantenuto i ghepardi come animali da caccia e come simbolo di prestigio sociale.) Secondo la cronaca scritta dal figlio, Akbar aveva fatto ricorso a gran parte delle sue enormi risorse per promuovere l'allevamento dei ghepardi e aveva persino dato a quegli esemplari regali la possibilità di scorrazzare per i giardini del suo palazzo. Ciononostante, fu partorita una sola cucciolata: fu questo l'unico evento del genere registrato nella storia dei ghepardi in cattività, fino a quando, nel 1956, nacquero dei cuccioli allo Zoo di Philadelphia.

Dal 1956 pochissimi programmi d'allevamento sono stati coronati da successo e solo il 10-15 per cento dei ghepardi sessualmente maturi, catturati allo stato selvatico, si è riprodotto in cattività. Questo basso indice di fecondità è spesso una conseguenza di malriusciti tentativi d'accoppiamento. Tuttavia, anche dopo accoppiamenti riusciti il ghepardo ha un indice di concepimento basso se lo si confronta con quello di altre specie allevate negli zoo; circa il 30 per cento dei cuccioli nati in cattività, inoltre, muore prima di raggiungere i sei mesi d'età.

Quando arrivammo al De Wildt per iniziare il lavoro di diagnosi, comprendente analisi dello sperma e uno studio sulla variabilità genetica, il centro possedeva circa 80 ghepardi. La colonia era stata impiantata con esemplari provenienti da due distinte regioni sudafricane: la Namibia (Africa sudoccidentale) e la regione settentrionale della provincia del Transvaal, nella Repubblica sudafricana. Queste regioni sono separate dal Deserto del Kalahari, cioè da una distanza di 1500 chilometri e quindi i responsabili del centro ritenevano che i vari animali utilizzati appartenessero a due sottospecie geografiche distinte.

La prima fase del nostro lavoro consistette nel raccogliere e analizzare 40 campioni di sperma provenienti da 18 maschi. Gli eiaculati erano molto diversi da quelli di altre specie che avevamo studiato. La

Due delle posizioni che il ghepardo, il più veloce scattista del mondo, assume mentre corre alla massima velocità. Le prove di velocità cronometrate dai ricercatori hanno fatto registrare valori massimi di 112 chilometri all'ora. Durante una battuta di caccia, il ghepardo resta sollevato da terra per la metà del tempo ed è in grado di raggiungere in pochi secondi la velocità di 80 chilometri all'ora con partenza da fermo.

concentrazione degli spermatozoi era solo un decimo rispetto a quella riscontrata normalmente nei gatti domestici e la mobilità era anch'essa significativamente inferiore. Ma, fatto ancora più sorprendente, la qualità degli spermatozoi era invariabilmente scadente. Nella media degli eiaculati, circa il 71 per cento degli spermatozoi aveva una forma anormale, in netto contrasto con il 29 per cento riscontrato nei gatti domestici. Per esempio i flagelli, organi propulsori degli spermatozoi, risultavano spesso arrotolati o ripiegati ad angolo retto, e molte teste erano o troppo grandi o troppo piccole.

I risultati di questi studi sullo sperma sembravano particolarmente significativi: infatti, la maggior parte delle anomalie riscontrate (particolarmente con un'incidenza così elevata) in altre specie è associata a sterilità. Per esempio, nel toro, il fatto che il 10-20 per cento degli spermatozoi sia anormale indica che l'animale ha una fertilità ridotta o sia addirittura sterile. Inoltre, è evidente che la morfologia degli spermatozoi è rigorosamente soggetta a controllo genetico; l'aumento di anomalie morfologiche appare spesso come una conseguenza di incroci tra consanguinei praticati tra animali di laboratorio o di allevamento. I dati sugli spermatozoi, raccolti al De Wildt, hanno fornito il primo indizio sulla natura del pericolo di estinzione che sta correndo il ghepardo.

Per valutare il grado di variabilità genetica nei ghepardi del De Wildt, abbiamo prelevato sangue da 50 esemplari e poi abbiamo estratto le proteine (per lo più enzimi) dalle cellule ematiche e dal plasma, e le abbiamo analizzate, nei campioni portati negli Stati Uniti, per mezzo dell'elettroforesi su gel, un metodo standard per lo studio della variabilità genetica. Per capire il senso di questo procedimento, si devono conoscere alcuni aspetti fondamentali dei meccanismi dell'eredità genetica.

Un animale riceve, al momento del concepimento, una doppia serie di cromosomi, una serie dalla madre e una serie analoga dal padre. Perciò ogni gene su un cromosoma possiede un gene corrispondente, detto allele, che è situato sul cromosoma appaiato al primo e svolge la stessa funzione, per esempio il controllo della sintesi di una proteina. Talvolta gli alleli ereditati da ciascun genitore e presenti in un determinato locus del cromosoma sono omozigoti, ossia identici; talaltra, invece, sono eterozigoti, ossia differenti. Un individuo che ha alleli eterozigoti per una data proteina produrrà due versioni di quella proteina.

Per effettuare un'analisi elettroforetica standard di campioni di sangue, i ricercatori isolano le proteine del sangue, le pongono su un gel e le espongono a un campo elettrico che provoca una migrazione delle proteine attraverso la matrice gelatinosa. Le proteine che sono esattamente uguali (prodotti di geni identici) migrano per distanze identiche nel gel. Le varietà di una determinata proteina (prodotti di geni che codificano per proteine aventi un amminoacido alterato nella loro impalcatura scheletrica) migrano, invece, su differenti posizioni. Le proteine possono poi essere messe in evidenza mediante un colorante specifico: perciò è possibile confrontarne la migrazione e determinare il grado di variabilità genetica nell'ambito di un singolo individuo e tra diversi individui (*si veda l'illustrazione a pagina 50*).

L'indagine elettroforetica sui ghepardi del De Wildt ha portato a un risultato sorprendente. Non si è registrata alcuna diversità nelle 52 proteine esaminate: all'esame elettroforetico ogni ghepardo appariva identico a tutti gli altri. Questo fenomeno è assai insolito. Analoghe indagini elettroforetiche sulla variabilità genetica, compiute su più di 250 specie nei vent'anni precedenti, avevano dimostrato che, in ciascuna specie, dal 10 al 60 per cento dei geni è polimorfico (ossia controlla la sintesi di più di una forma elettroforetica di proteina) e che dall'1 al 36 per cento dei loci genetici di un individuo medio è eterozigote. I geni dei ghepardi erano invece monomorfici (ossia controllavano la sintesi di una sola forma di proteina) e, quindi, era omozigote anche il locus per ciascuna proteina. Questo straordinario grado di monomorfismo genetico è raro nelle popolazioni naturali. Alla stessa stregua delle analisi compiute sullo sperma, esso richiamava alla mente l'omogeneità genetica osservata nelle specie inincrociate di proposito, per esempio certi topi di laboratorio.

L'albero evolutivo dei felidi indica le parentele tra alcune delle 37 specie della famiglia oggi esistenti. In base alla documentazione fossile, nel Miocene, tra 20 e 30 milioni di anni fa, apparvero parecchie famiglie di tigri dai denti a sciabola che più tardi si estinsero, con una varietà, *Smilodon*, giunta fino a 10 000 anni fa. I felidi moderni cominciarono la loro radiazione evolutiva dagli altri felidi all'incirca 12 milioni di anni fa. Il ramo antico portò ai felidi sudamericani di piccola e media taglia, tra i quali l'ocelot e il gatto tigre. Il secondo ramo si staccò circa da otto a 10 milioni di anni fa e portò al gatto domestico e ai suoi affini. Il terzo ramo, staccatosi circa quattro-sei milioni di anni fa, portò inizialmente a un ceppo di animali simili alle pantere (tra i quali i gatti dorati, il puma, il ghepardo e il serval) e successivamente a una biforcazione tra la lince e gli attuali felidi di grossa taglia. Questa filogenesi si basa su analisi effettuate su una proteina, l'albumina, nelle specie di felidi attuali.

L'attuale area di diffusione del ghepardo (*in colore*) viene posta a confronto con quella all'inizio del secolo (*in grigio*). Nel 1900 l'area di diffusione di *Acinonyx jubatus*, il ghepardo attuale, si estendeva per tutta l'Africa, il Medio Oriente e l'India. Oggi l'animale si è estinto dovunque come specie libera salvo che nell'Africa centrale e meridionale. Si stima che il numero attuale di ghepardi vari tra i 2000 e i 25 000 esemplari. Fossili di *A. jubatus* sono stati trovati in Europa, Asia, Africa e America Settentrionale. Sono stati trovati anche fossili di almeno tre altre specie di ghepardo estinte: *A. pardinensis*, ghepardo gigante che viveva forse quattro milioni di anni fa in Cina, India ed Europa meridionale; *A. intermedius*, specie più piccola che scorrazzava per tutta l'Eurasia, e *A. trumani*, che viveva negli Stati Uniti e presentava una certa rassomiglianza con il puma.

Nelle nostre ricerche sempre più raffinate per trarre indicazioni sulla variabilità genetica, abbiamo intrapreso un tipo un poco diverso di analisi delle proteine: l'elettroforesi bidimensionale su gel. Come la tecnica da noi adottata nelle ricerche sugli enzimi, questo procedimento si basa ugualmente sulla migrazione delle proteine in un gel, ma può analizzare simultaneamente centinaia di proteine. Abbiamo chiesto a David Goldman e a Carl R. Merril del National Institute of Mental Health di esaminare le proteine dei fibroblasti (cellule del tessuto connettivo) di sei ghepardi sudafricani non imparentati, viventi in zoo europei e statunitensi. Questi ghepardi presentavano sì alcune proteine varianti, ma la frequenza di polimorfismo riscontrata tra le 155 proteine analizzate risultava estremamente bassa: il 3 per cento, ossia meno di un terzo del valore che si osserva di solito nelle popolazioni umane studiate mediante la stessa tecnica.

L'immagine, che si veniva delineando, di una specie inincrociata che correva un insolito rischio genetico è stata rafforzata da tre altre ricerche, in primo luogo quella basata su un confronto tra il ghepardo e altri felidi. Andrea Newman, ricercatrice al laboratorio di uno di noi (O'Brien) al National Cancer Institute, ha eseguito un'indagine elettroforetica esauriente sugli enzimi delle cellule ematiche di nove specie di felidi, ospitate in zoo o riserve in tutto il mondo. I risultati hanno rivelato livelli di variabilità genetica da moderati ad alti all'interno di ciascuna specie studiata e tra le specie esaminate alcune avevano aree di diffusione sovrapposte a quella del ghepardo africano (leopardo, leone, serval e caracal), altre no (tigre, ocelot, gatto tigre e gatto domestico). In altre parole, i ghepardi da noi studiati erano eccezionali, anche nell'ambito delle specie di felidi affini, per il basso livello di variabilità genetica.

Il secondo insieme di dati è stato ottenuto da un'analisi dei caratteri morfologici effettuata su un gruppo di crani di ghepardo africano conservati in musei americani. Abbiamo valutato la variabilità morfologica sulla base dell'«asimmetria fluttuante», cioè l'entità della differenza di dimensioni in caratteristiche che in un dato individuo sono normalmente immagini speculari l'una dell'altra, come per esempio le parti destra e sinistra del cranio. In molte specie l'asimmetria aumenta di grado e di incidenza per effetto di accoppiamenti consanguinei, ma non è chiaro come insorga.

Robert K. Wayne, del nostro laboratorio al National Cancer Institute, ha calcolato l'asimmetria fluttuante in ciascuno dei 33 crani conservati nei musei e ha confrontato i risultati con misurazioni ricavate da leopardi, ocelot ed esemplari di *Felis tigrina*, specie che sapevamo (dai risultati della Newman) dotate di elevata variabilità genetica. I crani dei ghepardi presentavano un livello superiore di asimmetria rispetto ai crani delle tre altre specie di felidi.

Il terzo e più allarmante insieme di dati comprovanti il rischio genetico per il ghepardo emergeva da uno studio avente lo scopo di determinare se questa specie ha una qualche variabilità in corrispondenza

Spermatozoi aberranti scoperti in ghepardi del De Wildt Cheetah Breeding and Research Center sono posti a confronto con un esemplare normale (*fotografia a sinistra*). Tra le molte sorprendenti anomalie vi sono (*da sinistra a destra*) l'attorcigliamento del flagello, il ripiegamento del pezzo intermedio e la riduzione delle dimensioni della testa. La scoperta che i ghepardi del De Wildt avevano un'alta percentuale di spermatozoi anomali ha fornito il primo indizio che i problemi riproduttivi del ghepardo sono di natura fisiologica e possono avere una spiegazione genetica.

del maggior complesso di istocompatibilità (MHC). L'MHC è un locus genetico complesso, nei mammiferi, e dirige la sintesi di antigeni sulla superficie di gran parte delle cellule. Questi antigeni entrano in contatto con i linfociti *T* (globuli bianchi circolanti) durante una risposta immunitaria a un agente infettivo o a cellule aberranti che insorgono spontaneamente, come quelle tumorali. Gli antigeni codificati dall'MHC (antigeni MHC) sono anche responsabili del rigetto immunitario che segue ai trapianti di tessuti o di organi quando il donatore e il ricevente hanno diversi antigeni MHC. Negli ultimi 50 anni il sistema MHC è stato ampiamente studiato nell'uomo, nel topo, nel ratto, nel cane, nel cavallo, nello scimpanzé e in parecchie altre specie.

Il locus MHC è il più polimorfico dei loci dei mammiferi. Esso è in realtà un grappolo di subloci funzionalmente distinti, sul cromosoma, raggruppati nelle classi I, II e III. Nella specie umana il sublocus MHC appartenente alla classe I, da solo, è composto di tre geni, ciascuno dei quali può essere estratto da più di una dozzina di varietà alleliche presenti nel pool genico. Di conseguenza, in questa specie, il numero di possibili combinazioni nel sublocus della classe I, su un cromosoma, è enorme (più di 12^3) e la probabilità che esistano due individui che abbiano la medesima combinazione è minima (meno di una su 10 000). Anche altre specie di mammiferi presentano questo estremo polimorfismo. Le principali eccezioni sono costituite da ceppi di topi inincrociati, ottenuti intenzionalmente allo scopo di renderli omozigoti nei geni MHC; questi topi perciò accettano trapianti provenienti da membri dello stesso ceppo, non da altri ceppi di topi.

Per accertare l'esistenza di polimorfismo nel locus MHC dei ghepardi, abbiamo preso sette coppie di ghepardi dell'Africa meridionale: sei costituite da membri non affini e una costituita da fratelli. Tra i due membri di ogni coppia abbiamo scambiato trapianti di cute. Da ciascun animale è stato staccato un frammento di cute, che è stato poi tagliato a metà. Una delle due metà è stata suturata in sede adatta nel ricevente, che era l'altro partner della coppia; si trattava di un «allotrapianto». L'altra metà è stata suturata, invece, in sede adatta, nello stesso animale da cui era stato prelevato il frammento: si trattava di un «autotrapianto». Quindi gli animali sono stati esaminati e i loro trapianti sono stati ripuliti due volte alla settimana per circa otto settimane; durante tutto il periodo non è stato notato alcun segno di rigetto.

In parecchie specie di mammiferi, tra cui il gatto domestico, il tempo medio di sopravvivenza dei trapianti di cute provenienti da un animale non affine è di circa 10-12 giorni; il rigetto in questo periodo è da attribuirsi a differenze a livello di MHC. Il rigetto su tempi più lunghi, il rigetto cronico, è il risultato di differenze a livello di altri loci di istocompatibilità minori, quando, a livello del locus MHC, c'è identità. Se un ghepardo avesse un rigetto di un allotrapianto nell'arco di 10-12 giorni, si potrebbe concludere che i geni MHC dei due membri di una coppia non siano identici.

Ancora una volta, i ghepardi hanno dimostrato la loro straordinaria uniformità genetica: tutti gli allotrapianti sono stati accettati e anzi, per tutto il periodo di 10-12 giorni, non si sono potuti distinguere dagli autotrapianti. Tre allotrapianti hanno poi subito un lento rigetto, ma parecchi hanno resistito per almeno 78 giorni; passato questo periodo essi erano ormai confusi con la cute del ricevente.

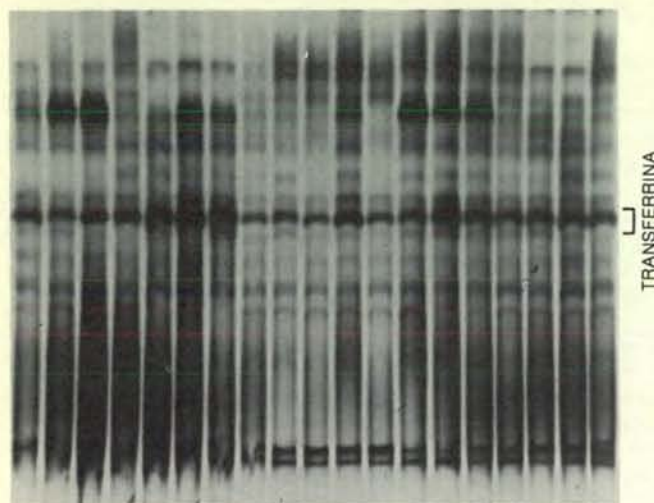
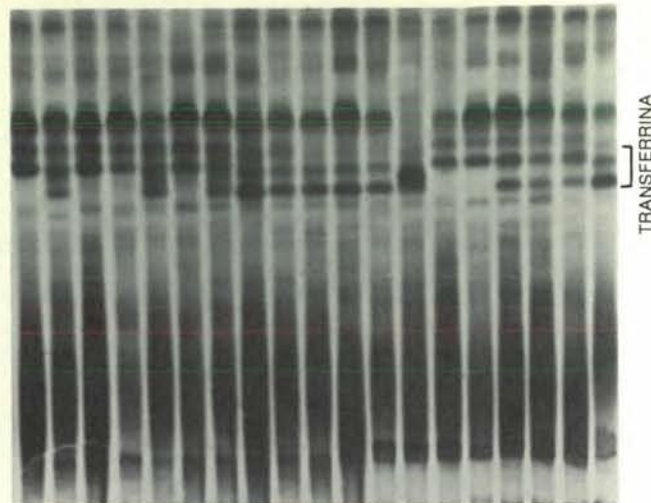
Per essere sicuri che l'assenza di rigetto degli allotrapianti fosse la conseguenza di una identità a livello dell'MHC e non di una incapacità generale del sistema immunitario del ghepardo a rigettare i trapianti di tessuti, abbiamo provato anche a suturare la cute di un gatto domestico in due ghepardi. In entrambi i casi il trapianto ha subito un classico rigetto di tipo acuto tra il decimo e il dodicesimo giorno, mentre sia l'autotrapianto sia l'allotrapianto continuavano ad attecchire e a crescere. Si è così dimostrato che il sistema immunitario del ghepardo era veramente in grado di riconoscere gli antigeni codificati da geni MHC estranei, mentre non aveva rilevato antigeni estranei nei trapianti cutanei provenienti da ghepardi presumibilmente non imparentati.

I ghepardi, in questi esperimenti di trapianto di cute (e, per questo argomento, in tutti gli studi che abbiamo descritto), dimostrano un livello di monomorfismo genetico che non ha precedenti in alcuna specie esoincrociata di mammifero. Tuttavia c'è una possibilità che va presa in considerazione prima di poter affermare definitivamente che i ghepardi di tutta l'Africa sono, in sostanza, gemelli genetici e prodotti di incroci consanguinei o inincroci. Gli studi biochimici considerati hanno interessato sempre campioni provenienti da ghepardi dell'Africa meridionale. La maggior parte dei crani che mostravano una maggiore asimmetria (indice di incroci consanguinei) proveniva, invece, da popolazioni di ghepardi dell'Africa orientale, ed era stata raccolta da Theodore Roosevelt e collaboratori. Vi è la possibilità che i ghepardi dell'Africa orientale e centrale, per quanto tra loro probabilmente consanguinei, abbiano pool genici in qualche modo diversi da quelli dei ghepardi sudafricani. Non nutriamo molto ottimismo su questo punto: tuttavia è un campo su cui stiamo ancora indagando.

Il fatto che l'uniformità genetica sia una minaccia per la sopravvivenza di una popolazione o di una specie era evidente già al tempo in cui Darwin formulò la teoria della selezione naturale. La variabilità genetica è la materia prima su cui interviene l'evoluzione; è l'eterogeneità genetica su cui opera la selezione naturale in periodi di cambiamento ecologico. La sorpresa consiste allora nel fatto che, nonostante tutto, il ghepardo monomorfico sia ancora nel novero delle specie viventi. Tuttavia, il pericolo di estinzione esiste. La teoria evoluzionistica prevede che una specie dotata di scarsa plasticità genetica sia particolarmente vulnerabile in un periodo di perturbazione ecologica e, nel caso del ghepardo, la dimostrazione si è avuta nel 1982.

Una coppia di ghepardi apparentemente sani giunse (in prestito per un programma di allevamento) al Wildlife Safari Park dell'Oregon che disponeva, per i ghepardi, di uno dei programmi meglio riusciti del mondo. Nel giro di qualche mese la coppia manifestò febbre, diarrea e itterizia e morì nello spazio di alcune settimane. La causa di morte venne diagnosticata come un'infezione virale: peritonite infettiva felina.

L'infezione, pur diffondendosi in modo virulento nelle colonie di gatti domestici, provoca di rado la morte di più del 10 per cento degli animali colpiti. Nel caso dei ghepardi, il fenomeno si è sviluppato in modo ben diverso. Entro sei mesi dai primi due decessi, sintomi di peritonite infettiva felina si manifestarono in tutti i ghepardi del Safari Park e, verso la fine del 1983, quasi la metà dei ghepardi era morta per questa malattia. Questa epizoozia (ossia «epidemia tra animali») è stata la risposta più grave a un'infezione di peritonite infettiva felina, finora registrata in una qualsiasi specie. Da allora siamo stati informati

Dei due metodi elettroforetici qui raffigurati e che servono a misurare il grado di variazione genetica nei ghepardi, uno (*in alto*) sottopone il sangue fresco all'eparina per impedire la coagulazione e per permettere di separare i diversi componenti (globuli bianchi, globuli rossi e plasma) mediante centrifugazione. Gli enzimi solubili estratti dai campioni di sangue vengono quindi sottoposti a elettroforesi, cioè vengono esposti a un campo elettrico che li fa migrare attraverso una matrice di gel, dopodiché sono resi visibili mediante coloranti specifici. Gli enzimi prodotti da geni polimorfici migrano in posizioni diverse nella matrice di gel, mentre gli enzimi prodotti da geni monomorfici migrano su posizioni identiche. Nel secondo metodo (*in basso*) le proteine marcate con elementi radioattivi ed estratte da cellule cutanee coltivate in vitro vengono esposte a campi elettrici che le separano secondo due dimensioni dello spazio. In primo luogo le proteine vengono separate in base alla carica elettrica e quindi in base al peso molecolare. Alla fine il gel viene esposto a una pellicola sensibile ai raggi X che rivela centinaia di proteine, le cui posizioni possono quindi essere messe a confronto per i vari ghepardi.

Le migrazioni elettroforetiche di una proteina, la transferrina, sono mostrate in 19 campioni di plasma estratti da ghepardi (*a sinistra*) e in altrettanti campioni estratti da gatti domestici (*a destra*). La transferrina sintetizzata dai ghepardi è migrata tutta in un'unica posizione (*banda scura in tutti i campioni*), indicando che i 19 ghepardi sono monomorfici, cioè portano un identico gene per la transferrina. Al contrario, la transferrina sintetizzata dai gatti domestici è migrata su diverse posizioni, rivelando la presenza di quattro diversi tipi e dimostrando che la popolazione dei gatti è polimorfica, cioè possiede più d'una variante per il gene responsabile della proteina. (Le lettere in basso indicano le combinazioni di due geni espresse da ciascun gatto.) Nei ghepardi sono stati trovati analoghi risultati uniformi, sostanzialmente per tutti gli enzimi studiati.

di analoghe epizoozie con identico agente eziologico in allevamenti di ghepardi in Irlanda, Canada e Namibia.

Una possibile spiegazione dell'epizoozia dell'Oregon è stata che il virus della peritonite infettiva felina apparteneva a un ceppo particolarmente virulento e che, per puro caso, aveva colpito per primi proprio i ghepardi. Quest'idea non ha però retto alle prove successive. I numerosi tentativi compiuti per trasmettere l'infezione a gatti domestici, inoculando in essi i virus estratti dai ghepardi infetti, non hanno portato ad alcun risultato. Inoltre, i dieci leoni africani, ospitati nello stesso parco, non hanno presentato alcun sintomo dopo essere stati esposti al virus. L'assenza di virulenza nei due gruppi di felidi ha dimostrato che, nei ghepardi, l'epizoozia era la conseguenza della vulnerabilità di questa specie e non dell'ipervirulenza del ceppo virale.

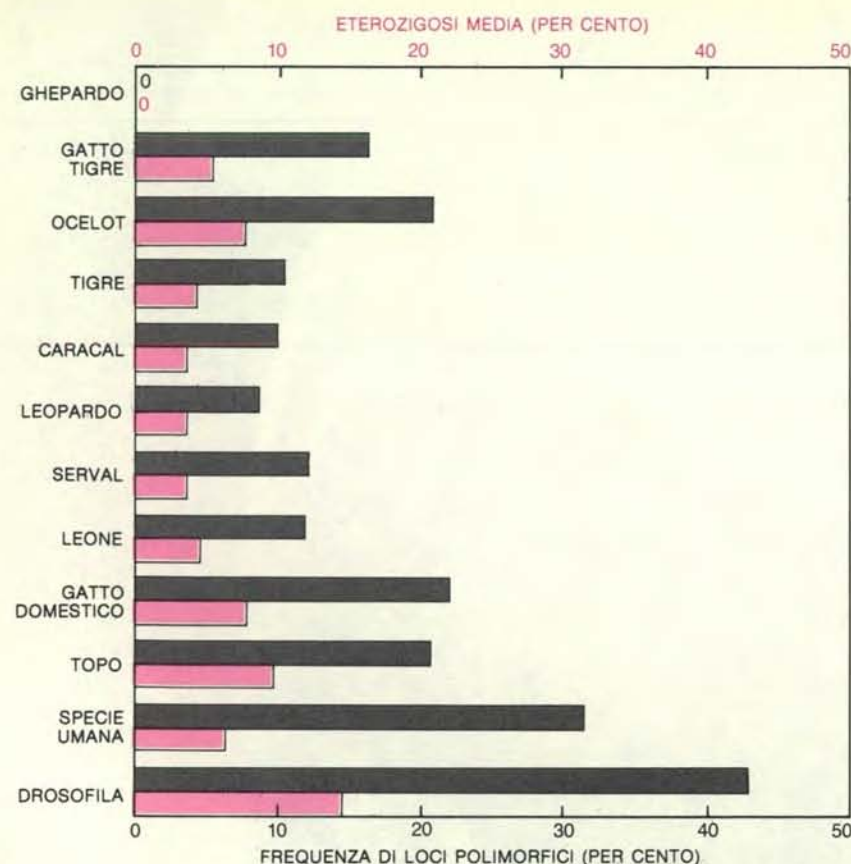
Il monomorfismo a livello del locus MHC, che limiterebbe il repertorio delle difese di una specie contro un virus, è l'indiziato numero uno per una spiegazione biologica dell'epizoozia. In una specie monomorfica sarebbe prevedibile una percentuale più elevata d'incidenza della malattia non appena l'agente virale ha sopraffatto le difese dell'organismo, perché tutti gli altri animali hanno più o meno lo stesso grado di suscettibilità al virus.

Il fatto avvenuto nell'Oregon potrebbe realmente essere spiegato grazie al monomorfismo dei geni nel sublocus MHC o della classe II o della classe I. Il sublocus della classe II contiene i geni responsabili della risposta immunitaria e si ritiene che contenga anche quelli della restrizione virale (i due tipi di geni potrebbero benissimo essere la stessa cosa). I geni della risposta immunitaria codificano per gli antigeni che stimolano, negli animali di laboratorio, la produzione di anticorpi contro antigeni sintetici o contro virus. È stato dimostrato che i geni della restrizione virale nei topi determinano il grado in cui un organismo può impedire la replicazione di virus il cui materiale genetico sia costituito da RNA; l'agente della peritonite infettiva felina è appunto un virus a RNA. Una popolazione che sia polimorfica in corrispondenza del locus MHC dovrebbe avere molte varietà di geni per la risposta immunitaria e per la restrizione virale con la capacità di conferire protezione contro il virus della peritonite infettiva felina. D'altra parte, una popolazione che divenisse improvvisamente monomorfica in corrispondenza del sublocus di classe II sarebbe particolarmente vulnerabile a un ceppo virale in grado di eludere le difese immunitarie controllate dall'insieme unico, condiviso da tutti, di geni di questa classe.

Il monomorfismo in corrispondenza del sublocus MHC di classe I produrrebbe una epizoozia con un altro meccanismo. I geni MHC di classe I codificano per gli antigeni che devono apparire sulla superficie di una cellula infettata da virus prima che i linfociti *T* possano riconoscere gli antigeni virali (i quali sono anch'essi esposti sulla superficie cellulare) e uccidere la cellula infettata. Un virus qualsiasi, in grado di alterarsi in modo da impedire che i linfociti *T* riconoscano, sulla cellula che ha infettato, l'antigene MHC di classe I, sarebbe molto efficace contro l'ospite. In una popolazione polimorfica, un simile adattamento a un ospite da parte di un virus non ha probabilità di essere efficace contro il sistema immunitario di un altro ospite, che abbia un diverso insieme di antigeni di classe I. Invece, in una popolazione monomorfica, un virus che si adatti al sistema di sorveglianza immunitaria di un animale troverebbe ogni altro sistema che incontra nella popolazione di quell'animale identico al primo e perciò facile da eludere.

Può essere importante ricordare che le strategie più sofisticate per eludere la sorveglianza immunitaria si evolvono nei virus in parallelo con i sistemi di difesa dei loro ospiti. Indipendentemente dall'esatta spiegazione della reazione della colonia di ghepardi al virus responsabile della peritonite infettiva felina, l'episodio avvenuto nell'Oregon sembra mettere in evidenza lo svantaggio in cui si trova una popolazione geneticamente depauperata nell'adattarsi agli agenti patogeni presenti nell'ambiente in cui essa vive.

Avendo scoperto tutta un insieme di prove che dimostrano che il ghepardo africano è una specie geneticamente monomorfica, ci siamo dedicati allo studio delle cause storiche del suo inincrocio che ha portato al monomorfismo. I dati che, nell'evoluzione dinamica di una specie, si riferiscono a un singolo punto nel tempo, non permettono di fare molti progressi. Ciononostante, abbiamo preso in esame parecchie ipotesi, ciascuna delle quali presume che gli antenati dell'attuale ghepardo avessero un tempo una variabilità genetica del tutto pari a quella che hanno oggi gli altri felini.

Secondo una di queste ipotesi, il ghepar-

Sulla base dei risultati di analisi elettroforetiche di enzimi, viene messa a confronto la variabilità genetica di 12 specie, tra cui nove felidi. (Loci eterozigoti sono quelle posizioni, sulle coppie di cromosomi, in cui i geni omologhi differiscono l'uno dall'altro.) Sulle oltre 250 specie studiate dagli esperti di genetica di popolazione, il ghepardo presenta il minimo valore di variabilità genetica.

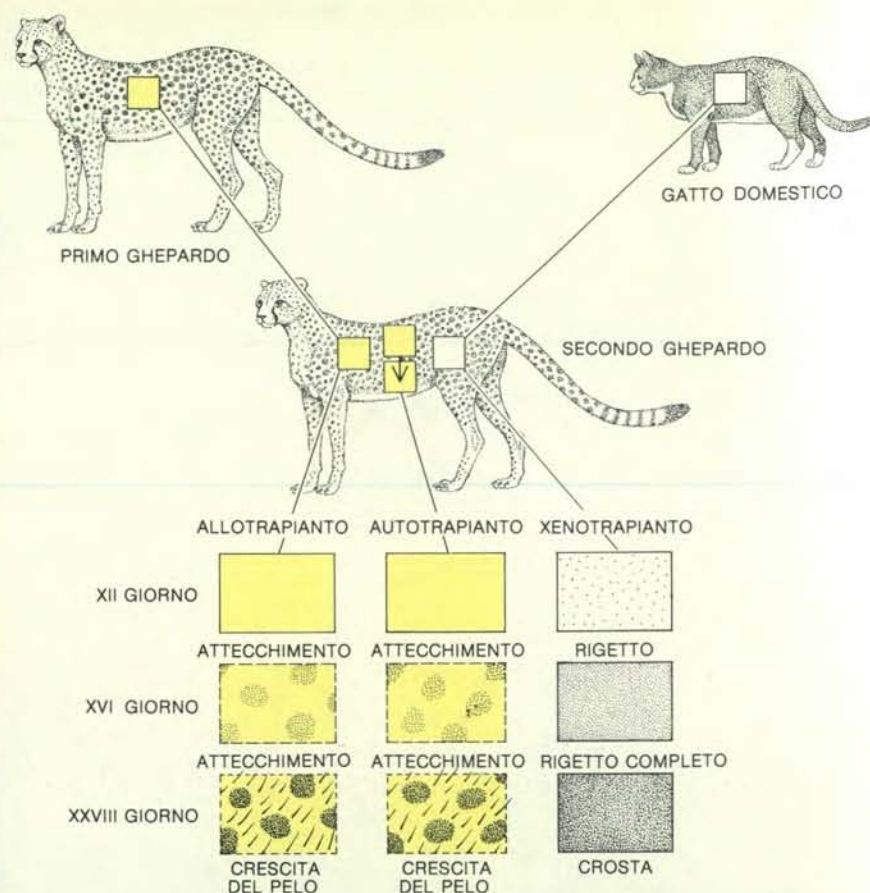
do avrebbe una predisposizione per gli accoppiamenti incestuosi, ma quest'idea non ha riscontro nelle osservazioni compiute dagli etologi che hanno studiato i ghepardi dell'Africa orientale e meridionale. Questi studiosi hanno rilevato che i ghepardi non rimangono in genere con i loro fratelli dopo l'infanzia. I maschi sono spesso animali territoriali e solo occasionalmente nomadi; femmine che non hanno alcuna parentela con essi vagabondano solitarie dentro e fuori i loro territori, accoppiandosi con parecchi maschi prima di spostarsi.

Per la seconda ipotesi il ghepardo si sarebbe evoluto fino a un optimum adattativo per una particolare nicchia ecologica e poi avrebbe perso gradatamente la variabilità durante un periodo prolungato di stabilità della nicchia. L'idea è particolarmente attraente per coloro che considerano il ghepardo un capolavoro di costruzione evolutiva, ma varie previsioni basate su di essa non sono confermate dai fatti. Per esempio, una specie adattata in modo ideale a una particolare nicchia ecologica non dovrebbe avere tra i contemporanei animali con stretti vincoli di parentela e, invece, i reperti fossili rivelano che almeno quattro specie di *Acinonyx* hanno scorrazzato sulla Terra negli ultimi milioni di anni. Si possono fare solo congetture sul numero delle altre specie esistite. Una specie adattata in modo ottimale dovrebbe competere con successo con altre specie nel suo habitat, ma l'attuale ghepardo ha scarse capacità di competizione: nella prova di forza con altri predatori, questi prevalgono sempre e i giovani ghepardi vengono spesso uccisi. Una specie adattata perfettamente alla sua nicchia dovrebbe essersi liberata gradualmente degli alleli recessivi deleteri e dovrebbe essere diventata omozigote solo per caratteri altamente adattativi: il ghepardo, invece, ha ancora una gran quantità di geni scadenti, per esempio quelli che controllano la morfologia degli spermatozoi.

Secondo la terza ipotesi, che è anche la più plausibile, a un certo punto nel passato la specie avrebbe attraversato un collo di bottiglia molto stretto: una assai grave riduzione numerica della popolazione. Il fenomeno sarebbe stato seguito, naturalmente, da una serie di inincroci che avrebbe diminuito il pool genico per la perdita casuale di alleli.

Quanto sarà stato stretto il collo di bottiglia per produrre una popolazione con una variabilità enzimatica dello zero per cento e una identità a livello del locus MHC? Quale è stata la causa di questa riduzione nelle popolazioni di ghepardi e quando il fenomeno si è realizzato? La teoria e la pratica dimostrano che una popolazione che passi attraverso un collo di bottiglia di soli sette individui può ancora mantenere circa il 95 per cento della variabilità genetica originaria: può mantenere questa variabilità se i sopravvissuti aumentano di numero rapidamente e secondo una progressione geometrica. (L'espansione lenta in una piccola popolazione aumenta la probabilità che scompaiano diversi tipi di geni.) Perciò sospettiamo che almeno una volta, e forse parecchie volte, nel passato le popolazioni degli antenati del ghepardo attuale abbiano avuto un crollo, riducendosi a un numero minimo di individui e sfuggendo per un pelo all'estinzione. Può anche darsi che i ghepardi sopravvissuti non siano riusciti a espandersi come numero rapidamente. Dio solo sa perché la popolazione di ghepardi sia diminuita così drasticamente. Le spiegazioni variano da una catastrofe climatica a una pestilenza di natura virale o batterica, alla distruzione dell'habitat o alla strage pura e semplice perpetrata dall'uomo.

Il momento in cui si è avuto il primo collo di bottiglia, come pure l'entità e la causa che l'ha determinato, è difficile da precisare. Secondo l'ipotesi più plausibile, un'importante riduzione numerica della popolazione si sarebbe avuta tra 10 000 e 12 000 anni fa, proprio alla fine del Pleistocene. Prima di questa data esistevano molte specie di *Acinonyx* e la specie attuale, *A. jubatus*, aveva un'area di distribuzione a livello mondiale. Poi qualche cosa provocò una massiccia estinzione di specie di mammiferi, particolarmente delle specie dell'America Settentrionale e Meridionale, distruggendone il 75 per cento. Molti grossi carnivori si estinsero e, tra essi, tutte le specie di ghepardo a eccezione di *A. jubatus*. L'area di distribuzione di quest'ultima specie si ridusse alla fine ai limiti attuali, cioè ad alcune zone dell'Africa. Non è stato affatto chiarito se la catastrofe ambientale del Pleistocene sia stata la vera causa della riduzione della popolazione. Secondo i dati in nostro possesso, non possiamo escludere la possibilità che una crisi più recente sia responsabile della vulnerabilità attuale del ghepardo: speriamo che i nostri studi, tuttora in corso, ci permettano di approfondire maggiormente la storia passata di questo animale.

Le nostre scoperte hanno rivelato molte cose sul passato e sullo stato attuale del ghepardo; ma un aspetto importante, che ci sembra non vada trascurato, è la valutazione delle prospettive per il futuro. L'uniformità genetica del ghepardo è certamente pericolosa per la sopravvivenza della specie, ma a nostro parere non dovrebbe essere interpretata come una sentenza di morte. E per tutta una serie di buone ragioni. Se il collo di bottiglia di cui abbiamo parlato ebbe luogo veramente nel Pleistocene, la selezione naturale, che impedisce la sopravvivenza degli individui con caratteri di grave disadattamento, ha sicuramente eliminato i geni più deleteri.

La prova del monomorfismo del ghepardo in corrispondenza del maggior complesso di istocompatibilità (MHC) si è avuta con tre tipi di trapianto cutaneo chirurgico: allotrapianti (provenienti da animali della stessa specie, non imparentati), autotrapianti (da sedi diverse di uno stesso animale) e xenotrapianti (da una specie diversa). Allotrapianti e autotrapianti sono attecchiti facilmente e al ventottesimo giorno dopo l'intervento hanno sviluppato le tipiche chiazze del mantello e pelo normale. Gli xenotrapianti, provenienti da un gatto domestico, sono stati invece rigettati il dodicesimo giorno. I geni MHC determinano il fatto che un trapianto venga accettato; escludendo la somministrazione di farmaci immunosoppressori, un trapianto sopravvive solo se il donatore ha gli stessi geni MHC del ricevente. Il fatto che i ghepardi abbiano accettato gli allotrapianti si può spiegare in due modi: o il sistema immunitario è stato soppresso o possiedono geni MHC identici. Poiché gli xenotrapianti hanno subito rigetto, è evidente che il sistema immunitario funziona correttamente: è quindi ovvio che sono identici i geni in corrispondenza del locus MHC.

Parecchie altre specie animali sono passate attraverso gravi riduzioni numeriche nella popolazione e sembrano oggi in ripresa: il ghepardo potrebbe avere la stessa buona sorte. L'elefante marino del Nord (*Mirounga angustirostris*) costituisce un esempio incoraggiante. All'inizio del secolo la sua popolazione era ridotta a circa 20 esemplari. In seguito all'approvazione (e forse a causa) delle leggi protezionistiche del 1922 e degli anni successivi, essa è aumentata al largo delle coste del Messico e della California. Oggi il numero di elefanti marini in quell'area raggiunge le decine di migliaia.

Infine, come abbiamo visto, la sottospecie di ghepardo dell'Africa orientale potrebbe essere geneticamente diversa dalla sottospecie dell'Africa meridionale. Se ciò fosse vero, i ghepardi dell'Africa orientale potrebbero essere introdotti nei programmi di allevamento degli animali in cattività: tali programmi, con poche eccezioni, si sono occupati in gran parte di ghepardi provenienti dall'Africa meridionale. Anche se i due gruppi sembrano essere, sulla base dei nostri metodi limitati, geneticamente simili, si potrebbe allevarli e farli riprodurre assieme. Ciò è stato tentato con successo allo Whipsnade Zoo, vicino a Londra. Secondo i rapporti, i ghepardi accoppiati in questo centro hanno prodotto cucciolate con una bassa mortalità.

Gli esseri umani hanno «civilizzato» per secoli l'ambiente selvaggio e il prezzo pagato è stato troppo spesso la degradazione delle piu prestigiose creazioni dell'evoluzione. Contemporaneamente, dallo studio degli animali sono state ricavate preziose informazioni per la biologia e l'evoluzione umane. Speriamo che in futuro queste conoscenze possano essere sfruttate per salvare le specie minacciate e in via di estinzione. Forse la scienza potrà anche aiutare il mammifero più veloce del mondo a vincere la corsa per la sopravvivenza.

La rinnovazione della pecceta subalpina

Il succedersi di generazioni della picea (abete rosso) nelle foreste di alta montagna dipende dal verificarsi di fenomeni difficilmente riproducibili mediante le tecniche selvicolturali e richiede periodi di alcuni decenni

di Pietro Piussi

Uno dei principali problemi della tecnica selvicolturale è costituito dalla rinnovazione del bosco. Alla fine di ogni ciclo produttivo gli alberi considerati economicamente maturi vengono abbattuti ed esboscati, e lo spazio reso libero viene occupato da piante più giovani. Queste nascono dal seme che cade dagli alberi adulti che crescono nelle vicinanze oppure possono venire introdotte dal selvicoltore mediante semine o piantagioni di piantine allevate in vivaio.

Da tempo ci si è resi conto dei vantaggi e degli svantaggi delle due forme di rinnovazione. Quella «artificiale» assicura un rapido ritorno della vegetazione arborea e una composizione specifica predeterminata; si può, per esempio, introdurre una specie assente nel popolamento adulto che cade al taglio. Quella «naturale» si insedia molto più lentamente, ma fornisce all'inizio un numero elevato di piante sulle quali il selvicoltore può operare una selezione; la maggiore densità assicura inoltre lo sviluppo di tronchi con caratteri tecnologici di pregio. Le piante nate spontaneamente subiscono una precoce mortalità dovuta alla concorrenza; risultano avvantaggiati gli individui che si insediano nelle stazioni più favorevoli e quindi si ha un migliore sfruttamento della potenzialità produttiva.

Più controverso è l'aspetto economico: i costi di impianto che si sostengono quando si opta per la rinnovazione artificiale possono essere ripagati da un accorciamento del tempo necessario a ricostituire il bosco, quindi da un ciclo produttivo più breve.

Sia per ottenere la rinnovazione naturale sia per praticare quella artificiale da qualche secolo ormai si sono sviluppate numerose tecniche particolari, adatte ai diversi tipi di bosco e ai diversi ambienti, sia fisici sia socioeconomici. Si tratta di tecniche basate in genere su analisi e interpretazione delle varie forme di utilizzazione, adottate in precedenza su basi empiriche, e dei risultati di vecchie piantagioni, eseguite in alcuni casi a scopo sperimentale.

La varietà di situazioni (ambiente fisico, tipo di bosco, forme di proprietà, manodopera, tipo di produzione richiesta) e i lunghi tempi necessari per ottenere l'insediamento del novellame (talvolta alcuni decenni) fanno sì che le conoscenze di base siano non di rado scarse e i risultati ottenuti siano difficilmente confrontabili.

Per l'ecologo la rinnovazione si presenta come un processo di successione che, nei boschi coltivati, si svolge sotto il controllo più o meno stretto dell'uomo. Questi rende disponibile, con il taglio degli alberi adulti, il terreno su cui si insedierà la nuova generazione di piante, assicura talvolta la disseminazione (o comunque la diffusione di nuove piante) e con tagli di diradamento regola i rapporti di concorrenza tra specie diverse o tra individui della stessa specie. In realtà il controllo che l'uomo svolge non sempre ha successo: i tagli di rinnovazione o gli impianti possono fallire e la ricostituzione del bosco risulta difficile. È più facile incorrere in insuccessi negli ambienti in cui i fattori ambientali sono particolarmente severi (aridità, mancanza di calore, parassiti), oppure quando le esigenze della specie arborea sono poco note.

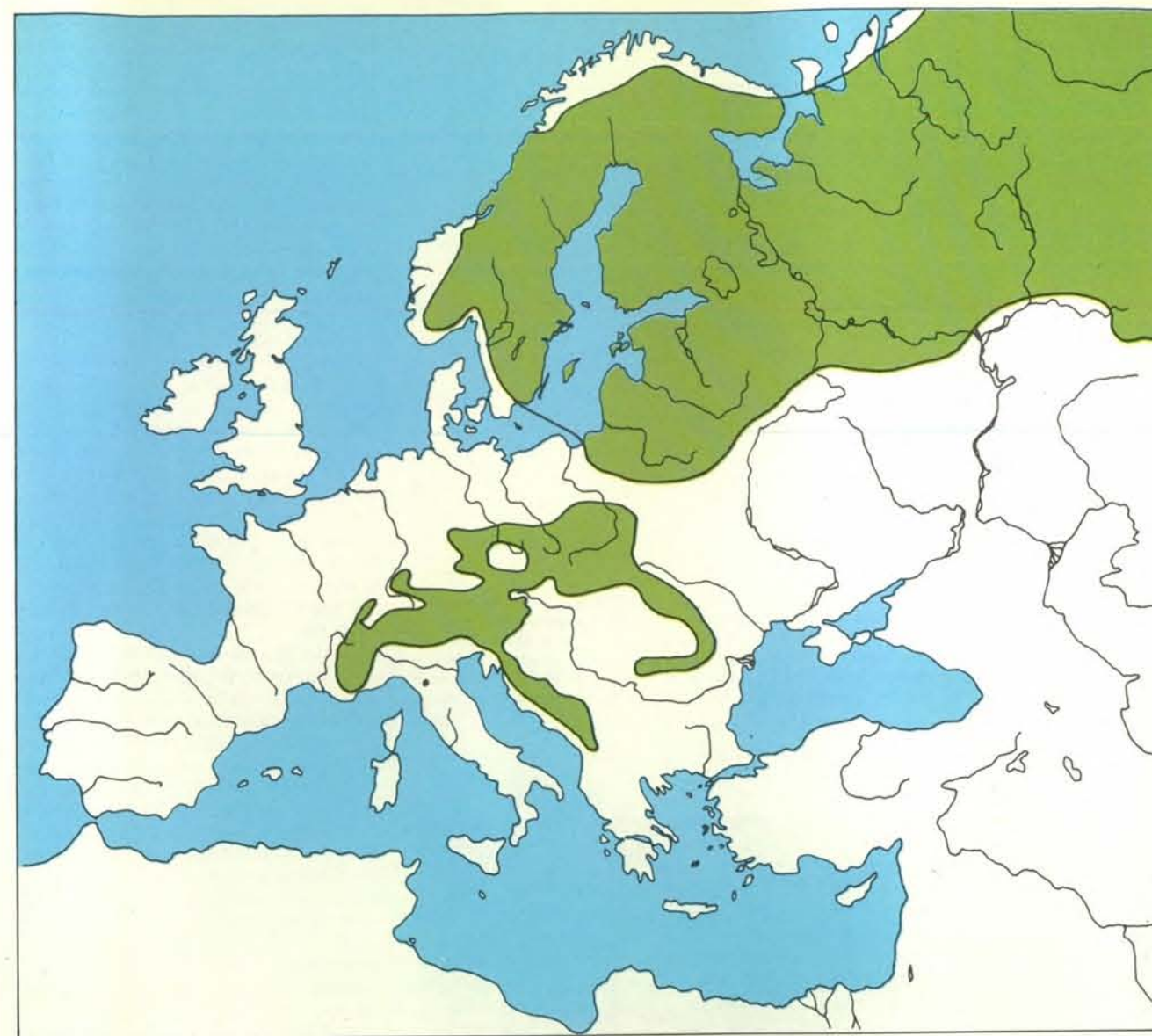
L'esame di queste situazioni, facendo uso delle conoscenze e dei concetti acquisiti nello studio delle successioni vegetali, può aiutare a comprendere le difficoltà e fornisce idee per sperimentare nuove soluzioni tecniche.

Un esempio molto interessante di questa problematica è costituito dalla rinnovazione della pecceta subalpina, formazione arborea dominata dalla picea che occupa la fascia più elevata del manto forestale sulle Alpi e su altre regioni montuose europee.

La picea o abete rosso (*Picea abies* Karst.) è una specie arborea che occupa in Europa un areale molto vasto: sull'arco alpino-dinarico, con una punta estrema sui monti Rodope nei Balcani, raggiunge i limiti meridionali della sua diffusione; è presente nell'Europa centro-orientale (nucleo carpatico-ercinico), ma occupa la maggior estensione in un nucleo boreale che dalla Scandinavia e dalla Prussia si estende verso oriente fino a toccare l'areale della specie asiatica affine *Picea obovata* (Ledebour), a volte ritenuta una sottospecie della *Picea abies*.

In questo ampio areale la picea occupa territori caratterizzati da climi molto diversi; può essere considerata una specie microterma che richiede tuttavia una certa quantità di calore durante il periodo estivo. D'altro canto, temperature estive troppo elevate e stazioni eccessivamente ricche d'acqua favoriscono altre specie arboree che esercitano una forte concorrenza sulla picea.

La picea è un albero di grandi dimensioni; può raggiungere i 50-60 metri di altezza, due metri di diametro alla base e un'età massima di oltre sei secoli. Anche nelle foreste italiane vi sono esemplari di taglia notevole, ma altezze di 35-40 metri e diametri superiori a un metro sono da considerarsi eccezionali.

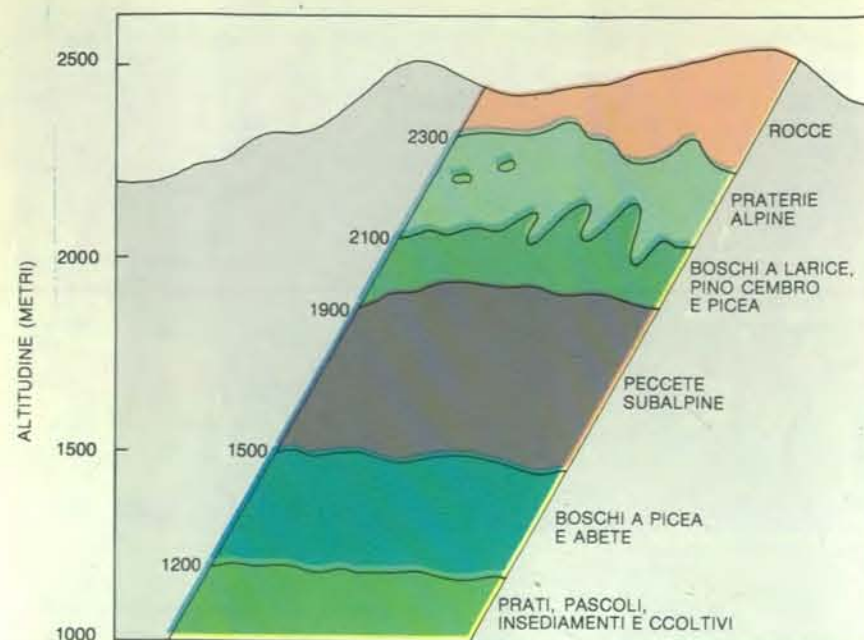
Le ottime caratteristiche del legno, che trova usi molteplici (legname da costruzione, per mobili, pasta da cellulosa per la carta, casse per strumenti musicali), la notevole plasticità ecologica, la rapidità di accrescimento giovanile e la facilità di coltivazione in numerosi ambienti, hanno fatto sì che da un lato i boschi di picea venissero pesantemente sfruttati da tempi remoti fino al punto da determinarne talvolta la distruzione, dall'altro che l'areale della specie venisse artificialmente allargato con piantagioni. In moltissimi casi, infine, la partecipazione di questa specie ai boschi misti è stata alterata profondamente in conseguenza di particolari tecniche selvicolturali e di altre forme d'uso del bosco (pascolo, raccolta della lettiera).

L'importanza economica della picea, o abete rosso (*Picea abies* Karst.), ha fatto sì che la sua piantagione sia stata diffusa su un territorio molto più vasto di quello originario. Nella cartina l'areale naturale della picea è raffigurato in verde. In Italia la specie è presente, oltre che sulle Alpi, in un piccolo tratto dell'Appennino settentrionale, alla Foce di Campolino, nella foresta dell'Abetone, la cui posizione è indicata dal pallino.

La picea è presente in Italia su circa 6000 chilometri quadrati in mescolanza con altre specie: nel piano montano, grosso modo al di sotto dei 1500 metri, tende a vegetare insieme all'abete bianco e, nelle zone con clima più oceanico, al faggio, ma l'azione dell'uomo ha favorito la formazione di boschi puri, anche a bassa altitudine. Nel piano subalpino forma in genere boschi puri, oppure costituisce, insieme al larice e al pino cembro, oltre i 1900-2000 metri, con una quota di partecipazione generalmente modesta, le formazioni del limite superiore del bosco che raggiungono eccezionalmente i 2200-2400 metri. Le peccete pure occupano in Italia circa 1300 chilometri quadrati.

In questi ultimi decenni le peccete subalpine sono state oggetto di studi approfonditi in diverse zone della Alpi. Le caratteristiche generali del clima, l'ecologia della specie e una certa omogeneità nelle forme di sfruttamento forestale del passato (dovuta a sua volta in larga misura all'uniformità dell'economia rurale delle valli alpine) permettono un utile confronto dei risultati ottenuti altrove e in qualche caso una loro generalizzazione.

Fino alla fine del Settecento nelle utilizzazioni boschive delle zone di alta montagna le esigenze economiche (fabbisogno di legname da opera e di combustibile oltre che di spazio per il pascolo degli animali domestici durante l'estate) hanno portato a trascurare i problemi selvicolturali. Solo nei primi decenni dell'Ottocento, in coincidenza con una intensificazione dello sfruttamento che riflette la riduzione delle risorse legnose disponibili nelle zone più accessibili e le accresciute esigenze della nascente industria, i tecnici forestali notano la difficoltà di ottenere la rinnovazione naturale, assai più facile nei boschi situati a minor altitudine e incontrano numerosi ostacoli nel supplire a tale mancanza con le piantagioni.

Vi è comunque ancora la tendenza ad applicare nell'ambiente alpino le conoscenze acquisite e i metodi usati per la picea nei boschi del piano montano, in un diverso contesto ambientale, quindi a considerare la specie e non l'ecosistema.

La soluzione dei problemi gestionali è poi ostacolata dalla situazione economica generale che impedisce di attuare chiare scelte tecniche: il pascolo in bosco è una pratica diffusissima, i terreni boschivi vengono messi a coltura temporanea o trasformati in pascoli, facendo non di rado un

La pendice volta a nord della Valle di Fiemme presenta una sequenza di piani vegetazionali determinati essenzialmente dalla variazione di temperatura, ma fortemente influenzati dall'azione dell'uomo. Nella parte inferiore si trovano gli insediamenti umani, i coltivi, i prati e i pascoli sfruttati in primavera e in autunno; più in alto vi sono i boschi a picea e abete, poi le peccete subalpine e infine i boschi a larice, pino cembro e picea. Oltre il limite della vegetazione forestale vi sono le praterie alpine usate come pascolo durante i mesi estivi e, più in alto ancora, le rocce.

All'interno di una peccata adulta il sottobosco copre quasi del tutto il suolo, da cui emergono le ceppaie di alberi abbattuti. I mirtilli, molto diffusi, sono un'importante fonte di alimento per l'avifauna della foresta. Le fotografie dell'articolo, esclusa quella di pagina 61 in alto, sono dell'autore.

impiego dannoso del fuoco per eliminare la vegetazione, i mezzi tecnici disponibili e la mancanza di strade per l'esbosco impongono tagli a raso che si estendono per decine di ettari.

Una intelligente analisi dei problemi di rinnovazione delle peccete era stata fatta alla metà dell'Ottocento da Joseph Wessely, funzionario forestale austroungarico. Tuttavia, a parte alcuni evidenti limiti nella conoscenza dei principi e dei processi biologici e a prescindere dai richiami «agli equilibri meravigliosi della natura», si nota che manca un modello di riferimento e quindi una interpretazione coerente in termini adattativi dei diversi fenomeni.

Verso la fine del secolo scorso si desta l'interesse per i boschi di montagna: una serie di catastrofiche alluvioni determina gravissimi danni alle cose e perdite di vite umane in numerose valli alpine. Ci si rendo conto della grande importanza della funzione protettiva esercitata dai boschi situati alle testate dei bacini fluviali: è qui infatti - su terreni generalmente ripidi - che si riscontrano le maggiori precipitazioni, che provocano piene dei torrenti, frane e valanghe.

Con il passare del tempo si perfezionano le indagini di dettaglio, i dati sperimentali, le osservazioni analitiche specialistiche, sempre alla ricerca di un fattore ambientale determinante, ma ci si allontana dalla visione unitaria del processo.

Più recentemente la rinnovazione è stata valutata nel quadro dell'evoluzione dell'ecosistema foresta e si è tentato di costruire modelli di funzionamento più generali in cui le attività forestali - tagli di utilizzazione e pascolo di animali domestici - vengono inquadrate come fenomeni di disturbo non diversi, sotto il profilo funzionale, da altri fenomeni naturali che interferiscono sulla struttura del sistema e sulla dinamica delle popolazioni.

L'opportunità di approfondire le conoscenze di base dei nostri ecosistemi forestali ha spinto a svolgere una descrizione accurata di alcune peccete situate sul versante meridionale delle Alpi, anche per poter compiere confronti più precisi con le peccete gravitanti nei settori settentrionali.

Le ricerche, rese possibili da contributi del Consiglio nazionale delle ricerche e dalla collaborazione della Provincia Autonoma di Trento, si sono svolte nella valle di Fiemme (in Trentino), che costituisce il tratto mediano della valle del fiume Avisio, affluente di sinistra dell'Adige, e in particolare nella foresta di Paneveggio la quale occupa, tra i 1450 e i 2200 metri sul livello del mare, il tratto superiore della valle del torrente Travignolo, principale tributario dell'Avisio.

Le valli dell'Avisio e del Travignolo hanno un decorso est-ovest e sono quindi protette dalla catena montuosa dei Lagorai che ne costituisce il versante sinistro e raggiunge altitudini di 2600-2700 metri sul livello del mare. Il clima risente quindi in parte di influssi continentali: le precipitazioni, nelle zone studiate, si aggirano intorno ai 1200-1600 millimetri annui con massimo estivo. In inverno cadono sotto forma solida e il manto nevoso ha una durata al suolo di quattro-sei mesi. La temperatura media annua va da tre a cinque gradi centigradi con un'escursione annua di 17-19 gradi. Occasionali gelate possono aversi anche in pieno luglio.

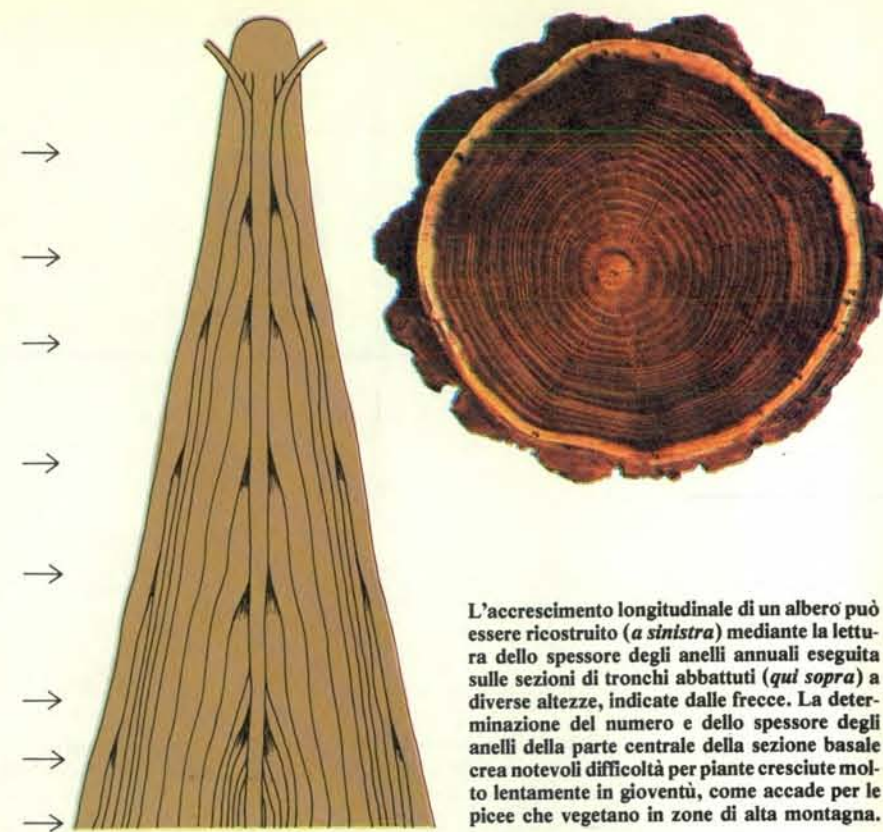
Nei popolamenti boschivi che possono essere definiti adulti, il tipo strutturale più diffuso è dato da un popolamento puro di picea al quale si possono mescolare rari larici o pini cembri. Le chiome dei singoli alberi sono strette, ma si sviluppano notevolmente in altezza: il fusto è così rivestito per oltre la metà della lunghezza, talvolta fino a qualche metro dal suolo, da rami vivi. Le altezze degli alberi non sono molto dissimili così che le chiome si riuniscono a formare un piano unico, spesso ma non compatto. Infatti anche se a volte due o più alberi crescono a stretto contatto, più frequentemente le chiome non si toccano, e una certa quantità di luce può filtrare fino al suolo. Questa luce consente l'esistenza di un sottobosco, spesso quasi continuo, di mirtilli, graminacee e muschi.

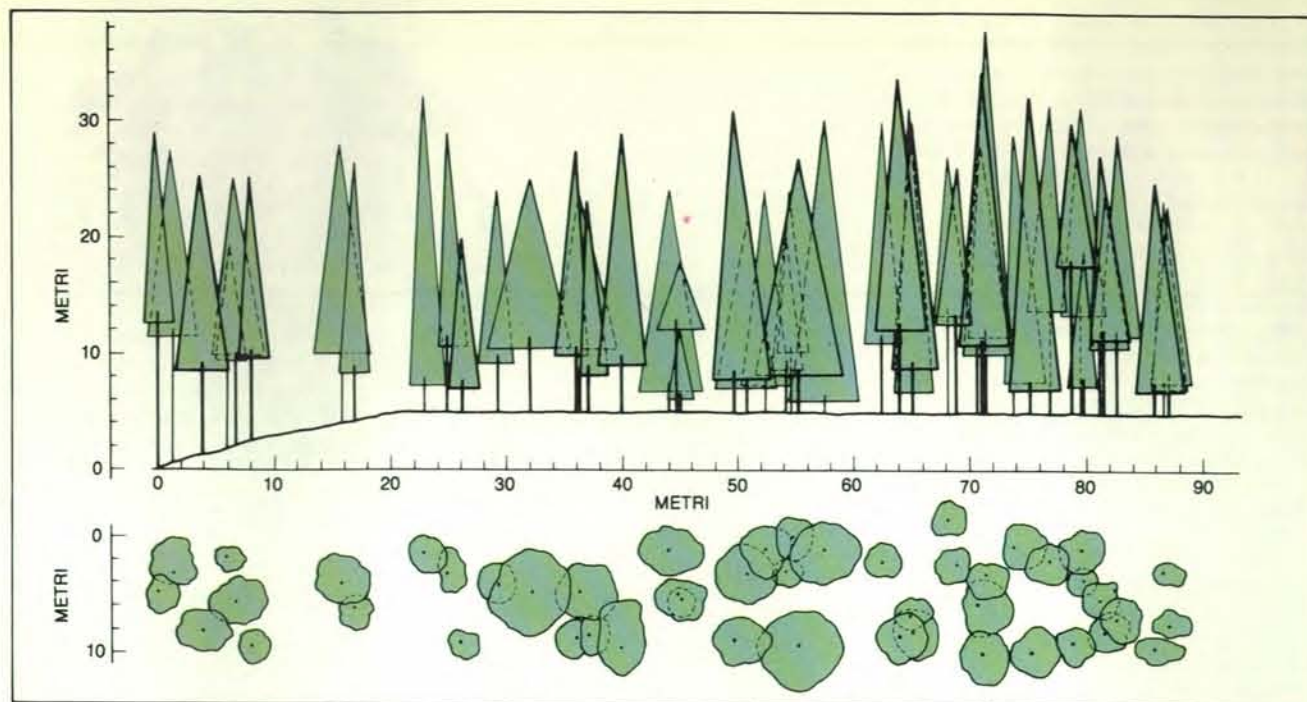
Frammiste alle erbe e agli arbusti del sottobosco vi sono altre piante di picea alte per lo più fino a due-tre metri, con densità inversamente proporzionale all'altezza. Questa seconda componente viene usualmente indicata come «prerinnovazione» in quanto si insedia prima che siano stati eseguiti dei tagli con cui si intende chiudere un ciclo produttivo e dare inizio a quello successivo.

Differenze di esposizione, di pendenza, di altitudine e di drenaggio del terreno o di disponibilità di sostanze nutritive determinano la presenza di varianti a questo quadro strutturale. Esso ha numerosi punti in comune con le peccete subalpine dell'Austria e della Svizzera. La struttura può essere ulteriormente definita con alcuni caratteri quantitativi: l'altezza delle piante più alte è compresa tra 25 e 35 metri, il diametro medio è di 30-40 centimetri (con massimi di 60-70 centimetri, eccezionalmente di più), il numero delle piante a ettaro va da 250 a 500, la massa legnosa da 300 a 800 metri cubi per ettaro.

Questa descrizione si riferisce a popolamenti adulti, molto diffusi, nei quali l'accrescimento, sia in altezza sia in diametro, è ridotto, la mortalità è minima e di conseguenza la struttura del popolamento si mantiene quasi invariata per periodi di qualche decennio.

Le indagini palinologiche svolte nella zona rivelano che, dopo una fase iniziale nella quale il dominio era tenuto dai pini, la picea si è affermata con continuità fino a oggi. Ricerche archeologiche segnalano una lunga presenza di cacciatori mesolitici (intorno al 6000 a.C.); non si può escludere che essi abbiano fatto uso del fuoco per cacciare la selvaggina o per ridurre il bosco a pascolo e abbiano in questo modo favorito la diffusione dell'ontano verde che appare in misura non trascurabile nei depositi pollinici.

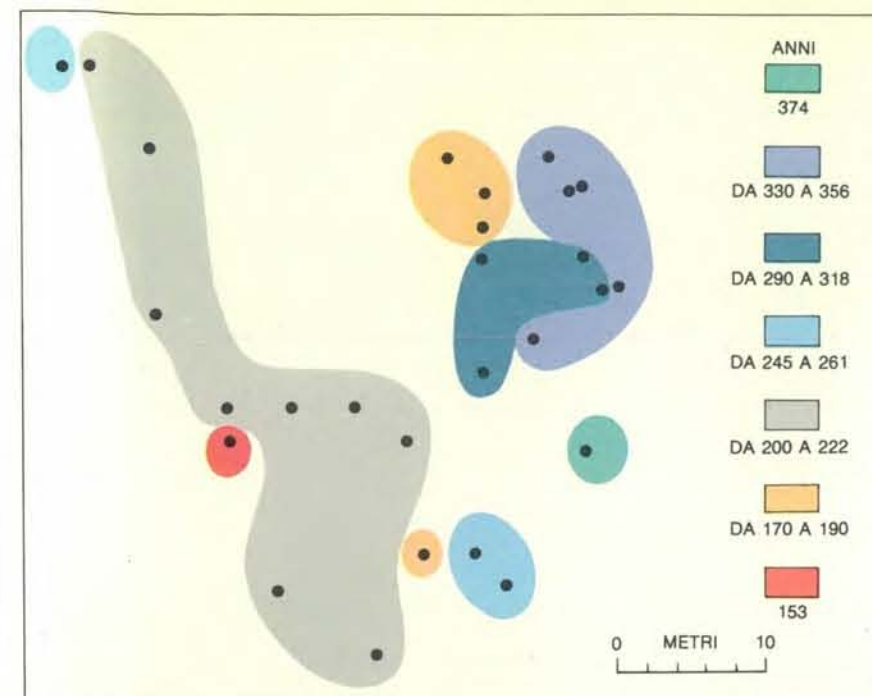
Le utilizzazioni di legname sono docu-

L'accrescimento longitudinale di un albero può essere ricostruito (*a sinistra*) mediante la lettura dello spessore degli anelli annuali eseguita sulle sezioni di tronchi abbattuti (*qui sopra*) a diverse altezze, indicate dalle frecce. La determinazione del numero e dello spessore degli anelli della parte centrale della sezione basale crea notevoli difficoltà per piante cresciute molto lentamente in gioventù, come accade per le picee che vegetano in zone di alta montagna.

Questa picea insediatasi in una chiaria del bosco, benché non più alta di un metro, conta quasi sicuramente oltre mezzo secolo di vita. Il portamento della pianta denota frequenti lesioni ai rami e alla cima, causati molto probabilmente dal peso della neve. I rami più bassi, che presentano numerosi aghi brunastri, sono stati danneggiati da funghi che si sviluppano sotto la neve.

La struttura di una pecceta adulta fa sì che le chiome raccolte intorno al fusto lascino filtrare la luce al suolo in misura sufficiente a consentire lo sviluppo di un piano arbustivo ed erbaceo quasi continuo. Anche la neve non trattenuta dagli alberi si accumula al suolo dove si scioglie all'inizio dell'estate. Il disegno riguarda la foresta di Paneveggio. Nonostante la grande uniformità di struttura, le piante, o piccoli gruppi di piante, hanno età molto diverse anche se poco distanziate.

La produzione di strobili è più elevata nelle piante di picea che hanno una chioma ampia e libera ai lati, mentre le picee, la cui chioma è soffocata dagli alberi circostanti, hanno fruttificazione scarsa o nulla. Nella fotografia, una chioma di picea durante un anno di abbondante fruttificazione. Gli strobili si concentrano nella parte alta della chioma e sono più numerosi nel settore esposto a sud.

mentate solo da qualche secolo, ed è raramente possibile ricostruire dalle carte di archivio la storia dei singoli tratti di bosco. Sembra comunque che solo tra la fine del Settecento e i primi decenni dell'Ottocento si sia praticato il taglio raso nelle foreste di alta quota, già in uso da molto tempo nelle zone più a valle. A Paneveggio le prime utilizzazioni massicce sarebbero state eseguite nei primi decenni dell'Ottocento; verso la metà del secolo circa un quarto della superficie della foresta era stato sottoposto a tagli rasi. Già allora gli amministratori si erano resi conto della difficoltà di ottenere la rinnovazione. Da quell'epoca, e per oltre un secolo, la storia della foresta è caratterizzata dai tentativi di assicurare la ricostituzione del bosco per via naturale o artificiale. La prima guerra mondiale e le vicende politico-amministrative di questi territori hanno ostacolato queste iniziative oppure hanno determinato la perdita della documentazione relativa. Oltre che per la produzione legnosa, la foresta di Paneveggio, così come tutte le foreste di montagna, è stata utilizzata per il pascolo di animali domestici. Solo in questi ultimi anni il carico di bestiame si è sensibilmente ridotto e vari settori della foresta non vengono più sfruttati per attività pastorali.

Il punto di partenza per le indagini era quindi dato da boschi di origine almeno in parte naturale, ma sicuramente modificati nella densità e nella struttura da una plurisecolare azione antropica.

L'attenzione si è concentrata su alcuni popolamenti adulti che, data la documentazione esistente, erano da ritenere sicuramente di origine naturale. Si è esaminata anzitutto la distribuzione degli individui in classi di età. Nonostante le chiome raccolte a una stessa altezza facessero pensare a un insieme coetaneo di alberi, è apparso invece che piante di dimensioni quasi uguali, situate a distanza di pochi metri tra di loro, avevano età molto diverse, talvolta fino a due secoli; le piante più vecchie superavano i tre secoli d'età.

L'analisi del fusto, compiuta su sezioni del tronco prelevate a diverse altezze, ha consentito di capire in quale modo le singole piante si sono accresciute. Durante la prima parte della vita questi alberi hanno avuto un accrescimento longitudinale molto ridotto: uno-due centimetri all'inizio, poi tre-quattro centimetri all'anno. Questo periodo è stato di durata diversa per i singoli individui: alcuni di essi avevano «atteso» 10 o 20 anni, altri 100 o più, prima di manifestare un accrescimento normale. Tuttavia, una volta superati i due-quattro metri di altezza, l'accrescimento longitudinale è progressivamente e fortemente aumentato, raggiungendo i 30-35 centimetri all'anno, per poi ridursi a 10-15 centimetri quando l'albero ha raggiunto i 20-25 metri di altezza. Le difficili condizioni di vita durante il periodo iniziale non influiscono quindi sulla longevità e sulle dimensioni raggiunte dagli alberi di picea.

Si poteva in tal modo supporre che alberi cresciuti sotto la copertura della chioma di alberi più grandi per oltre un secolo si comportassero come piante giovani una volta che fosse stata eliminata la concorrenza dall'alto. Questo fatto è espressione della differenza che i biologi rilevano tra una «età cronologica», espressa dal numero di anni che intercorrono dalla nascita, e una «età fisiologica» assai più significativa, che riflette la fase vitale in cui l'albero si trova.

Una verifica di questa ipotesi è stata fatta estendendo l'indagine alle piante di picea indicate come «prerinnovazione».

La distribuzione di queste piante è molto irregolare, sia se consideriamo tutta la foresta, sia se ci limitiamo a un solo soprassuolo. La prerinnovazione è in genere più abbondante sui versanti a solatio. Inoltre, in corrispondenza dei margini, dove il bosco adulto confina con popolamenti più giovani di bassa statura, essa è più densa quando il margine stesso riceve direttamente la radiazione solare (margini esposti a sud o margini «caldi»).

Nell'interno del bosco la densità registra i valori più bassi, anche se non nulli, e la distribuzione delle piante sul terreno è estremamente irregolare. Infatti le piantine si possono più facilmente incontrare dove vi sono particolari condizioni di substrato: legno marcio - vecchie ceppaie o residui di utilizzazione - e orizzonti profondi (*B* e *C*) del profilo del suolo portati alla luce da una accidentale rottura dell'orizzonte organico di superficie. Ciò può essere dovuto al trascinamento di un tronco, a uno smottamento di terreno, ma più frequentemente allo sradicamento accidentale di un albero adulto. Questo fenomeno determina, in effetti, la presenza di entrambi i substrati prediletti dalla picea. Su questo aspetto del problema ritorneremo in seguito.

Si deve notare che la picea è, in pratica, la sola specie arborea rappresentata. Rarissime sono le piante giovani di larice, di pino cembro e di sorbo degli uccellatori, ma sono anche molto rare in queste peccete le piante adulte di altre specie in grado di disseminare.

Su superfici piccolissime di terreno si trovano piante di età diversa: individui di due-tre anni accanto ad altri di 30 o 40. La densità è scarsa così che non vengono mai a crearsi condizioni di concorrenza tra le piantine.

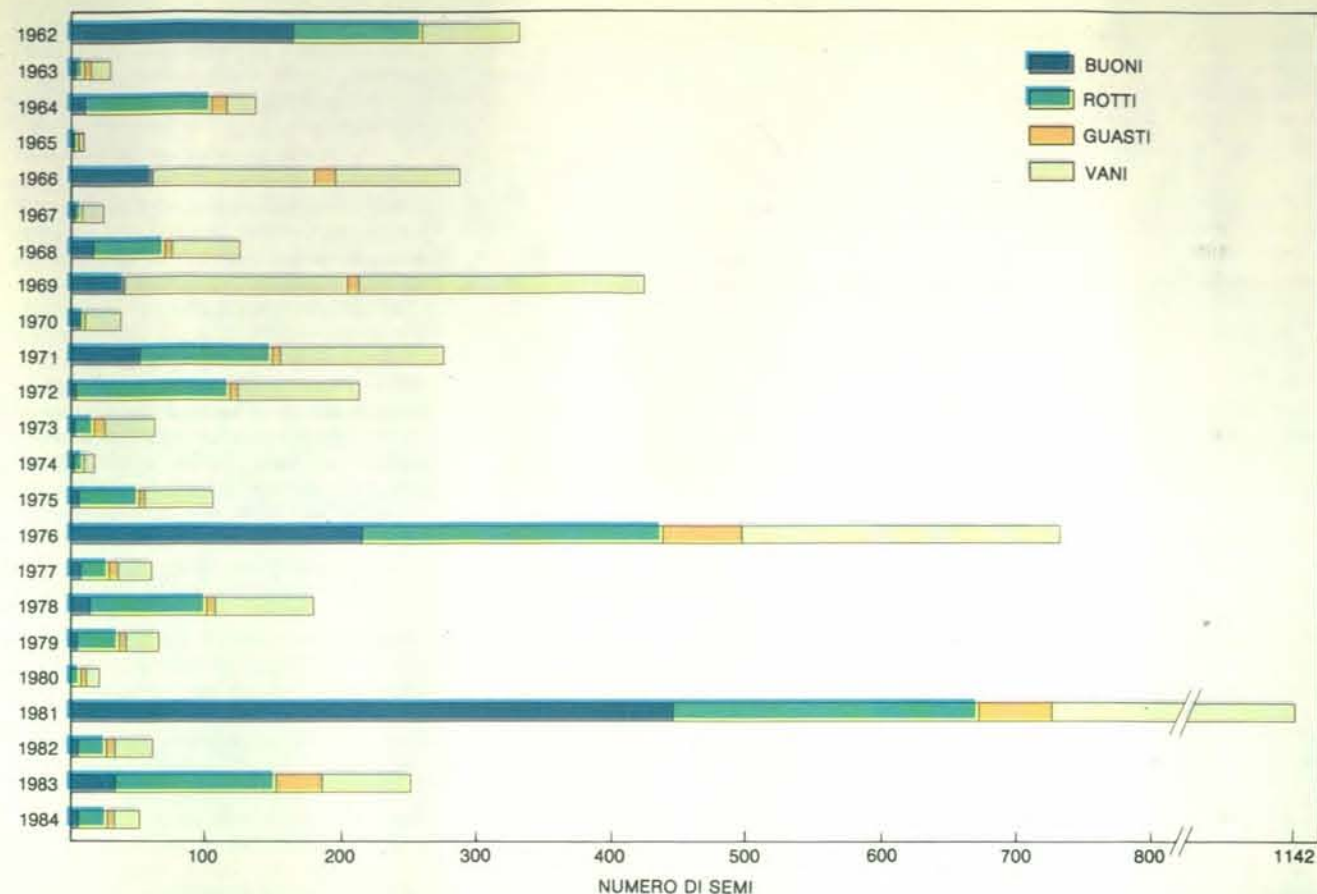
L'accrescimento è molto modesto: uno o due centimetri all'anno in media. La struttura delle età non è regolare: alcune classi cronologiche sono assai più rappresentate di altre. La prerinnovazione è quindi formata da individui di età diversa, insediatisi con irregolarità nel tempo e nello spazio, talvolta da molti anni.

Le caratteristiche di questo frammento «giovane» della popolazione - la disetaneità e la lentezza di accrescimento - ricordano quelle della prima fase di vita degli alberi adulti quali ci appaiono dalla ricostruzione fattane tramite l'analisi dei fusti; in altri termini si può dire che le modalità di insediamento attuali non sembrano differire da quelle di 200-300 anni fa. Il novellame, anche se apparentemente scarso, sarebbe sufficiente a ricostituire la copertura arborea allorché la vecchia generazione morisse o venisse abbattuta.

L'irregolarità dell'insediamento da un anno all'altro può essere spiegata con le caratteristiche di fruttificazione della picea. È noto che la maggior parte delle specie forestali non fruttifica con costanza annuale: le annate di produzione abbondante sono distanziate tra loro di due-tre o più anni. Le cause di questo fenomeno non sono completamente chiare e vi sono alcune interpretazioni che non si escludono vicendevolmente quali l'impossibilità di disporre ogni anno delle grandi quantità di sostanze necessarie alla produzione di seme, le condizioni favorevoli del tempo, una strategia adattativa per impedire la coevoluzione di animali predatori che renderebbero impossibile la sopravvivenza del seme e quindi la conservazione della specie.

La descrizione dei popolamenti adulti e della prerinnovazione aveva fornito una prima immagine degli aspetti più significativi della biologia della picea nei riguardi della rinnovazione: la sequenza temporale (produzione di seme, nascita, accrescimento iniziale e mortalità) e le differenze delle condizioni di vita esistenti tra stazioni diverse o, nella stessa stazione, in momenti diversi.

Dai dati raccolti emergeva - cosa abbastanza ovvia - che la disponibilità di seme, la radiazione solare e le condizioni del substrato influivano sulla densità della rinnovazione: era possibile però individuare solo delle tendenze e non stabilire, sia pure orientativamente, una scala assoluta per valutare la mortalità della popolazione nelle diverse condizioni.

È stato possibile ottenere maggiori informazioni dallo studio della produzione di seme disponendo nella foresta apposite

Dai dati sulla produzione di seme nella foresta di Paneveggio emerge che solo una parte, talvolta esigua, del seme prodotto è in grado di germinare, mentre è spesso elevata la percentuale di seme distrutto dagli animali. Gli anni di produzione abbondante non sono frequenti e sono irregolari.

«trappole» per raccogliere il seme nel periodo della disseminazione.

La fioritura della picea avviene tra giugno e luglio e il seme matura verso la fine di settembre. La schiusura degli strobili, che consente la fuoriuscita del seme, ha luogo a partire da novembre nelle giornate di sole e di vento. Il seme, piccolo e munito di un'ampia ala, può cadere ad alcune decine di metri, talvolta oltre, dalla pianta madre. Nella pecceta subalpina il seme cade in buona parte sulla neve e quindi raggiunge il suolo al momento dello scioglimento, ossia tra aprile e giugno. Al crescere dell'altitudine gli strobili si sviluppano di meno in lunghezza, la quantità di seme in essi contenuta diminuisce e la percentuale di seme germinabile si riduce.

La produzione di seme oscilla notevolmente nel lungo periodo; in 22 anni di osservazione si sono avuti solo tre casi di abbondante produzione, con oltre 100 semi in grado di germinare per ogni metro quadrato. Di tutto il seme prodotto in 22 anni, il 57 per cento è caduto in quattro anni. Le irregolarità nella struttura delle età della prerinnovazione e nella sostanziale disetaneità dei soprassuoli trovano quindi una spiegazione più precisa. Con le trappole è stato anche possibile quantificare la predazione. La quantità del seme spezzato dagli animali sfiora il 35 per cento del totale del seme prodotto e, dato che il seme spezzato era verosimilmente pieno e sano, diventa il 60 per cento del totale del seme inizialmente in grado di germinare. Molto più modesta è la frazione di seme danneggiato da funghi e da larve di insetti.

I predatori più importanti sembrano essere gli scoiattoli e gli uccelli (soprattutto il crociere) fintanto che il seme è contenuto nello strobilo, e i topi quando il seme è al suolo.

La predazione non costituisce pertanto un fattore trascurabile; nonostante ciò, nelle annate di buona produzione, rimangono sul terreno da 150 a 400 semi buoni per metro quadrato. La produzione di seme può sembrare scarsa in assoluto, ma in realtà è necessario valutare il modo in cui le saltuarie annate di piena produzione vengono messe a frutto.

Il secondo aspetto sul quale era necessario approfondire l'indagine era quello della nascita, dell'accrescimento iniziale, della mortalità e dei fattori che intervenivano in questi eventi. A tale scopo si sono eseguiti numerosi esperimenti in laboratorio, in vivaio e in foresta.

Eliminando la variabile costituita dagli animali predatori di seme mediante opportune protezioni delle semine e concentrando l'attenzione sui fattori fisici, si è notato che l'humus grezzo è un substrato assai più sfavorevole del sottostante terreno minerale per la nascita e l'accrescimento delle giovani piante di picea; il fattore determinante è risultato essere l'acqua, sia per la germinazione del seme e la nascita delle piantine sia per il loro accrescimento successivo. Il rifornimento d'acqua trova un ostacolo nella struttura fisica dell'humus grezzo; il disseccamento in superficie è facile, mentre è assai difficile la risalita d'acqua dal basso una volta che l'orizzonte olorganico si è disseccato.

L'umidità viene quindi assicurata e la germinazione è elevata quando il seme si trova all'ombra; è sufficiente un piccolo ostacolo fisico, quale un ramo secco caduto al suolo, per creare le condizioni adatte alla nascita. Se invece l'ombra è esercitata da ostacoli più estesi, come la vegetazione erbacea e arbustiva, oppure le chiome degli alberi adulti, nel giro di alcune settimane la riduzione di luce si fa sentire sulle plantule e ne limita notevolmente la nascita e la sopravvivenza.

È difficile separare, negli esperimenti svolti in bosco, l'effetto della luce da quello del calore e tenere distinti, per quest'ultimo, gli effetti positivi (processo di assimilazione) da quelli negativi (maggiore consumo d'acqua, scottature al colletto).

Nelle prime settimane di vita, le piantine, il cui fusto è ancora erbaceo e la radichetta poco sviluppata, reagiscono positivamente alla luce, ma in misura maggiore se nate su terreno minerale. Qui esse fruiscono di un migliore rifornimento d'acqua e così evitano l'effetto negativo del calore che si accompagna alla forte insolazione. Un effetto analogo si nota nelle piantine nate su humus grezzo quando queste vengono innaffiate.

La localizzazione preferenziale delle giovani piantine su terreno minerale può essere così interpretata come la più adatta ad assicurare livelli tollerabili di luce e di acqua. Qui sarebbe anche assicurata una migliore nutrizione minerale.

Le condizioni più sfavorevoli si hanno quindi, e ciò sembra un paradosso, sull'orizzonte umifero che è il substrato maggiormente diffuso nel bosco, mentre quelle più favorevoli si hanno sugli orizzonti minerali che costuiscono degli «accidenti», frutto di ciò che può apparire un disturbo alla struttura regolare del bosco. Il substrato minerale è sempre più o meno localizzato e occupa una piccola frazione del terreno.

Condizioni favorevoli si hanno pure sul legno marcio, un altro substrato apparentemente «accidentale». Nelle foreste da lungo tempo sottoposte a utilizzazioni esso occupa una superficie piccolissima ed è rappresentato da ceppaie di alberi abbattuti, da tronchi abbandonati dopo il taglio perché guasti e totalmente inutilizzabili e da altri residui di piccole dimensioni delle utilizzazioni boschive. Rare sono le piante morte per cause naturali e crollate al suolo. Queste sono invece evidentemente più numerose nei boschi vergini oppure in quelli sottoposti a utilizzazioni irregolari o abbandonati da lungo tempo: in essi il legno marcio è un substrato di nascita largamente diffuso.

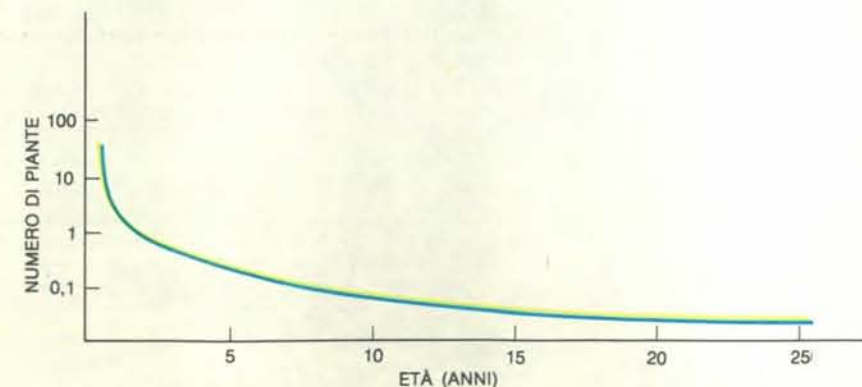
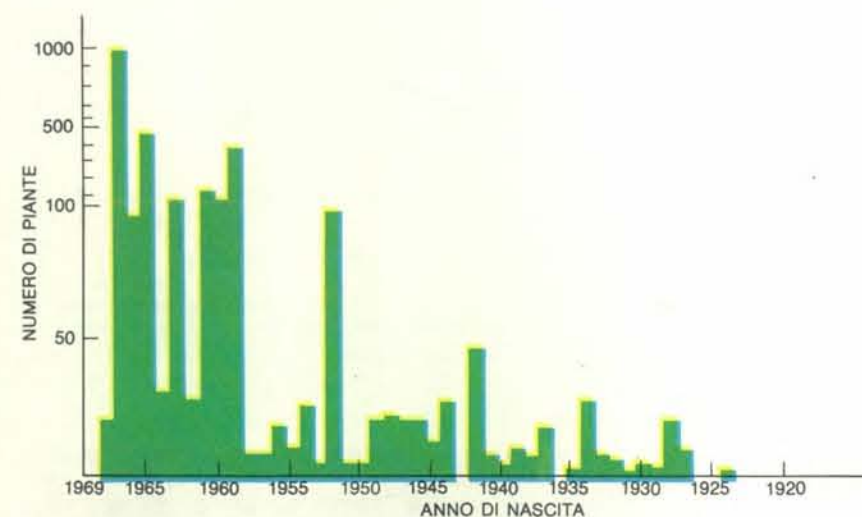
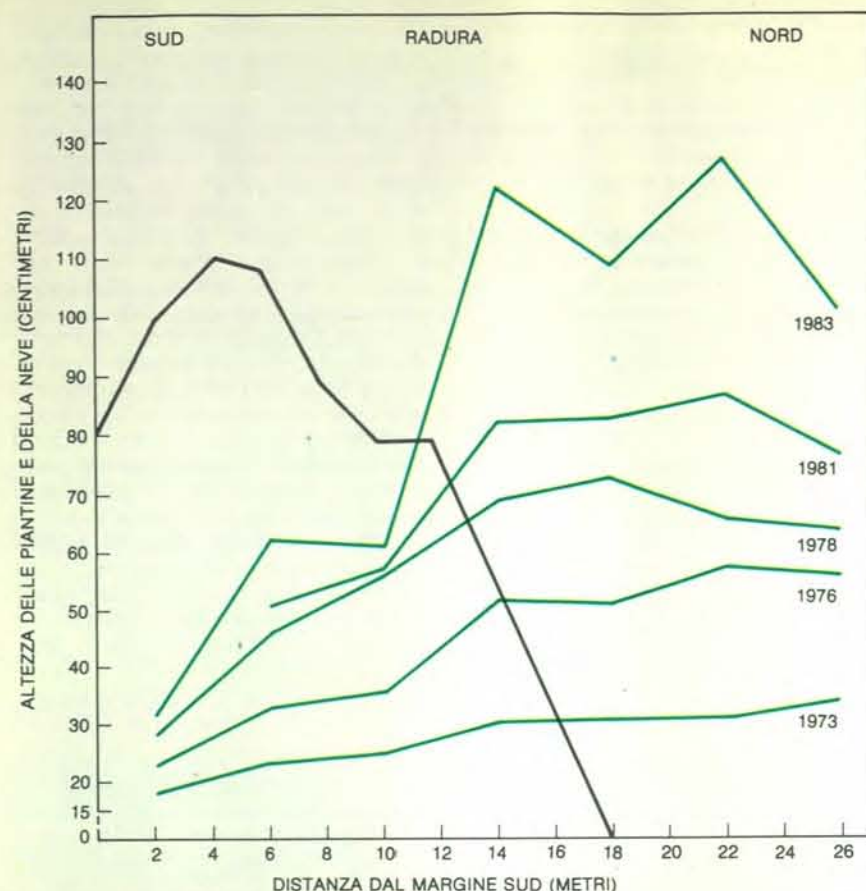
È piuttosto facile vedere piantine insediate sulle ceppaie e sui tronchi marcescenti e abbiamo anche osservato nei popolamenti studiati che le piante adulte sono spesso radicate su questo substrato. Il 45 per cento di esse ingloba alla base del fusto grossi frammenti di legno marcio. Su questo substrato le giovani piante si trovano in una situazione favorevole sotto vari aspetti: migliore rifornimento idrico, minore concorrenza di erbe e arbusti, minore pressione da parte della neve e più breve durata del periodo di innevamento e quindi anche un periodo di vegetazione più lungo.

L'esperimento più complesso è stato condotto in due piccole chiarie create con il taglio di tutti gli alberi (taglio raso) su una superficie di circa 1500 metri quadrati ciascuna. In queste tagliate si sono create delle condizioni di microclima analoghe a quelle che si avrebbero dopo il crollo di un gruppo di 50-60 piante adulte. Con una lavorazione localizzata del terreno è stato preparato in vari punti un substrato abbastanza simile a quello che, in caso di sradicamento, sarebbe stato disponibile per i semi.

Schianti e sradicamenti non sono fenomeni che si verificano con regolarità. Essi sono causati principalmente dalla caduta di neve primaverile pesante, oppure da venti molto forti. Anche la superficie interessata dai crolli di alberi è molto variabile: sono frequenti crolli isolati, ma si possono verificare distruzioni su superfici più estese: ettari o decine di ettari.

Entro la chiaria creata con il taglio e nel bosco circostante sono stati eseguiti semine e impianti; questi ultimi hanno permesso di valutare con maggiore dettaglio la situazione in cui si trovano piante alte 20-25 centimetri, che hanno già superato la fase particolarmente critica dei primi anni che seguono la nascita. Sono state esaminate anche alcune piante di prerinnovazione sparse irregolarmente entro la tagliata e si è seguito l'insediamento di nuove piante di disseminazione naturale.

I risultati confermano l'interpretazione precedentemente formulata sulla base delle indagini di struttura e di accrescimento.

Per quanto riguarda le semine, la natalità è stata massima nelle zone ombreggiate del bosco adulto circostante la tagliata; qui però si è avuta anche la massima mortalità durante il primo anno e l'accrescimento delle piante sopravvissute è stato scarso. Le condizioni ottimali di accrescimento si sono verificate nelle microstazioni più calde e luminose. In queste, le migliori condizioni per la sopravvivenza e per l'accrescimento si sono notate nella parte centrale della chiaria, mentre il margine ombreggiato è stato nettamente sfavorevole. Osservazioni specifiche hanno dimostrato che qui vi sono temperature del suolo e dell'aria più basse che in corrispondenza del margine soleggiato, più lunga permanenza della neve al suolo, minore luminosità, forti attacchi agli aghi da parte di funghi parassiti criofili. La concorrenza radicale da parte delle piante adulte si fa evidentemente sentire lungo tutto l'orlo della tagliata, ma può essere evidenziata solo lungo il margine soleggiato dove sono assenti o deboli altri fattori avversi. U-

La popolazione di giovani piante di picea che si insedia sotto i popolamenti adulti è formata da piante di età molto diversa che testimoniano la gradualità dell'insediamento. La frequenza degli individui per ogni classe di età (*in alto*) è proporzionale alla quantità di seme prodotto (sono evidenti i picchi in corrispondenza delle annate di produzione abbondante) e alla mortalità. È stato possibile delineare una curva di sopravvivenza delle piante di picea nella pecceta subalpina (*in basso*) ponendo a confronto i dati di produzione di seme con il numero di piantine presenti. La mortalità è massima nel corso del primo anno e le maggiori perdite si hanno verosimilmente nel seme non ancora germinato e nelle plantule erbacee durante il periodo che va dallo scioglimento della neve alla fine di luglio; le perdite nei primi 10-15 anni di vita non sono tuttavia trascurabili.

L'altezza media in anni diversi di giovani picee di uguale età e altezza iniziale, piantate nel 1970 in una radura creata con il taglio del bosco adulto, è indicata dalle curve in colore. Le condizioni lungo il margine freddo della radura sono nettamente sfavorevoli, a causa anche della lunga permanenza al suolo della neve, il cui spessore è stato rilevato il 10 maggio 1971 lungo un allineamento (*curva in nero*). Nella radura, infatti, lo scioglimento della neve procede progressivamente dal lato soleggiato a quello che si trova in ombra. Con il passare del tempo, anche in prossimità del margine caldo si osserva una riduzione dell'accrescimento degli alberi, molto probabilmente a causa della concorrenza esercitata dalle radici di quelli che crescono lungo il margine.

Novellame di picea si è insediato sul terreno sconvolto dall'apparato radicale di un albero adulto che è stato abbattuto dal peso della neve. Una parte delle giovani piante è morta, molto probabilmente in seguito al disseccamento del terreno nel corso di una estate particolarmente siccitosa.

guale reazione hanno mostrato le piante di prerinnovazione; con esse è stato anche possibile dimostrare come piante lungamente aduggiate siano in grado di riprendersi quando ricevono una sufficiente illuminazione.

Una pianta che si insedia spontaneamente impiega molti anni prima di sfuggire all'azione dei fattori che determinano la mortalità giovanile. La sua sopravvivenza dipende da numerosi fattori ambientali, alcuni dei quali agiscono costantemente o con una ciclicità di breve periodo, altri invece si manifestano con notevole irregolarità e scarsa frequenza. Di conseguenza, non solo la loro individuazione a posteriori è molto difficile e la loro interazione con altri fattori diventa in pratica impossibile da percepire, ma, a priori, la loro azione è indefinibile e non consente la costruzione di un modello accurato di coltivazione.

Ciò non è valido solo per una foresta vergine, ma anche per una foresta coltivata in cui la lunghezza dei cicli e la difficoltà di operare massicce trasformazioni sulla comunità e sull'ambiente fisico lasciano ampi spazi all'azione dei processi naturali. Di conseguenza, la grande difficoltà incontrata nel valutare l'effetto di fattori che agiscono con grande irregolarità spaziale e temporale rappresenta una seria limitazione; infatti i fattori che si manifestano in modo casuale - vento e neve soprattutto - agiscono positivamente nei riguardi della rinnovazione.

La conservazione della copertura arborea in un bosco, che si realizza con la sequenza di processi da noi indicata come rinnovazione, richiede che al momento della morte di uno o più alberi, allorché si crea la disponibilità di un tratto di terreno, si verifichino anche condizioni d'ambiente, sia climatico sia edafico, che consentano un facile insediamento, un buon accrescimento e una ridotta mortalità delle nuove generazioni di piante.

Si può così capire perché le condizioni dell'ambiente create dal crollo di alberi siano quelle più adatte all'insediamento di nuove piante.

Se, allo stato attuale delle conoscenze, non è possibile definire l'andamento della rinnovazione in maniera precisa, si può tuttavia tentare una migliore valutazione dei dati disponibili relativi all'ecologia delle peccete, anche in relazione agli effetti pratici della gestione del bosco.

Si è visto che le condizioni necessarie per la presenza della rinnovazione - disponibilità di seme ed esistenza di condizioni ambientali particolari - non si riscontrano con frequenza. Ciò è vero, tuttavia, solo se si effettua un confronto con altri tipi di bosco, anche di picea, situati in condizioni ambientali più favorevoli. Una valutazione di frequenza (ripetizione nel tempo e nello spazio di una particolare situazione) su basi biologiche e non economiche dovrebbe essere invece fondata sulle caratteristiche del sistema a cui si fa riferimento. Le osservazioni in foresta dimostrano che, in passato, la rinnovazione si è insediata entro tempi, per noi, molto lunghi. Il problema del tempo apparirà più chiaro se colleghiamo la natalità alla mortalità naturale; o, meglio, se lo inseriamo in un modello di evoluzione naturale dei boschi alpini con ampia partecipazione di specie sciafile. Ogni popolamento arboreo tende a saturare lo spazio aereo disponibile e gli individui che soccombono alla concorrenza lasciano nel piano delle chiome alcuni spazi che vengono utilizzati dai sopravvissuti più vigorosi. Questo processo avviene in popolamenti relativamente giovani formati da individui vigorosi. In seguito, su piante più vecchie, la mortalità, determinata soprattutto dall'azione di parassiti e da fattori meteorologici, diventa più forte e gli spazi lasciati liberi non possono più venire occupati dai sopravvissuti che hanno già raggiunto la taglia massima e la cui vitalità è in declino. Si creano quindi i presupposti ambientali per l'insediamento di una nuova generazione che può vegetare sotto una copertura più o meno densa determinata dalla vecchia generazione.

La graduale scomparsa del popolamento adulto che viene sostituito da alberi giovani sempre più numerosi, vigorosi e alti, può avvenire gradualmente, ma, data la labilità strutturale della vecchia generazione, può anche verificarsi con un collasso brusco provocato da fattori esogeni.

Nel caso della pecceta subalpina, dove altre specie arboree sono praticamente assenti, sembra legittimo supporre l'esistenza di un processo di autosuccessione, a meno che non si verifichino tagli rasi molto estesi o crolli altrettanto vasti, che rendono possibile l'insediamento, accanto alla picea, del larice.

Le osservazioni relative alla pecceta subalpina si adattano a questo schema generale; più precisamente, sembra di poter delineare due meccanismi di rinnovazione. Il primo consiste nello sfruttamento di condizioni ambientali particolari - abbondanza di luce e substrato costituito da terreno minerale e da legno marcio - per una colonizzazione relativamente rapida di un tratto di terreno privato improvvisamente della copertura arborea. Si può prevedere una elevata percentuale di nascite su piccoli tratti di terreno, ciascuno di qualche metro quadrato, compresi entro la superficie più grande. Il successo di questo meccanismo è maggiore se l'insorgere di queste condizioni coincide con una annata di abbondante produzione di seme.

Il secondo meccanismo consiste nel trarre profitto dalla capacità della picea di tollerare l'ombra nel periodo giovanile per accumulare gradualmente una scorta di nuove piante che potranno sostituire gli individui della generazione precedente quando questi muoiono. In questo secondo caso la percentuale di nascite è bassa, ma l'insediamento avviene, con tempi molto lunghi, su tutta la superficie del terreno. Questa modalità risponde all'esigenza di sostituire singoli individui che muoiono senza provocare forti sconvolgimenti dell'ambiente; è il caso di piante uccise da attacchi parassitari o anche stroncate da vento o neve senza che avvenga uno sradicamento che provoca lo sconvolgimento del terreno.

Una situazione del primo tipo non esclude che, prima dell'evento catastrofico, vi sia stato un insediamento di tipo graduale. Le soluzioni possibili sono anche legate alla struttura del soprassuolo e alle condizioni generali dell'ambiente edafico e climatico.

Le difficoltà tecniche che si presentano al selvicoltore deriverebbero quindi in buona parte dall'aver elaborato una tipologia dei boschi di picea eccessivamente semplice, dall'essersi proposto tempi troppo brevi per l'ingresso della rinnovazione e dall'aver attuato i tagli di maturità secondo schemi spaziali che non sono aderenti alla realtà naturale.

La tecnica selvicolturale nel corso di questi anni recenti ha tratto profitto dall'esperienza passata; anche se il substrato adatto potrà essere creato solo parzialmente con apposite lavorazioni del terreno che sono molto costose, una parte delle difficoltà sarà risolta riconoscendo che sono necessari periodi di rinnovazione più lunghi e impiegando forme di taglio che permettano di valorizzare meglio la prerinnovazione naturale.

Lo sradicamento di alberi adulti è un fenomeno non infrequente nei boschi e assume una importanza particolare dove l'uomo non interviene con tagli. La modifica delle condizioni di microclima, di suolo e di topografia favorisce l'insediamento di numerosi semenzali sul terreno smosso e sul legno marcio. Nel disegno sono rappresentate tre fasi successive a distanza di 10-20 anni l'una dall'altra. Lo sradicamento di un albero allontana la copertura di humus grezzo (*in verde*) e l'orizzonte lisciviato (*in grigio*) e mette allo scoperto l'orizzonte di accumulo (*in marrone chiaro*).

Lo stesso nucleo galattico è visto con un telescopio ottico (*in alto*) e con una schiera di radiotelescopi sui quali sono montati rilevatori a effetto tunnel a superconduttori (*in basso*). L'immagine radio a falsi colori è una carta dell'emissione delle molecole di monossido di carbonio presenti nella galassia e corrisponde al rettangolo in negativo dell'illustrazione in alto. La fotografia nel visibile è stata realizzata al Palomar Observatory e l'immagine radio all'interferometro per onde millimetriche di Owens Valley del California Institute of Technology. La galassia ripresa è M82.

I rilevatori a effetto tunnel in radioastronomia

Questi dispositivi a semiconduttori sono molto sensibili e consentono agli astronomi di esplorare regioni dell'universo inaccessibili ai telescopi ottici, come per esempio le nubi di gas interstellare dove si formano nuove stelle

di Thomas G. Phillips e David B. Rutledge

L'astronomia è spesso una lotta incessante da parte dell'osservatore per raccogliere una quantità sufficiente della debole radiazione elettromagnetica che proviene dalle stelle lontane al fine di compiere misurazioni significative. Per aiutarsi in questa lotta i ricercatori hanno seguito due strategie: hanno costruito telescopi via via più grandi in modo da raccogliere una maggior quantità di onde elettromagnetiche e hanno aumentato la sensibilità alle onde captate dei rilevatori montati sui telescopi. La costruzione di giganteschi telescopi ottici e radio può essere molto costosa. Fortunatamente, un raddoppio della sensibilità di un rilevatore può avere altrettanta influenza sull'efficienza di un telescopio di un raddopppio dell'area di apertura dello strumento. Per questo motivo gli astronomi sono sempre stati particolarmente desiderosi di migliorare la sensibilità dei rilevatori.

Lo sforzo ha portato negli ultimi dieci anni alla realizzazione di uno strumento chiamato rilevatore a effetto tunnel a superconduttori. A parte la sua sensibilità straordinariamente elevata, questo rilevatore è importante anche perché è eccezionalmente adatto alle bande millimetrica e submillimetrica, ossia le gamme di lunghezze d'onda in cui opera la radioastronomia. Più esattamente, il campo di attività della radioàstronomia riguarda la misurazione di radiazioni celesti aventi lunghezze d'onda comprese tra varie migliaia di metri e pochi decimi di millimetro (la regione radio dello spettro elettromagnetico). Con gran parte dei metodi attualmente in uso è difficile rilevare radiazioni all'estremità della banda verso le piccole lunghezze d'onda.

Agli astronomi piacerebbe avere un maggior numero di informazioni su queste regioni per due motivi. Innanzitutto perché i luoghi di formazione di nuove stelle si trovano nelle profondità di nubi di gas interstellare e quindi non sono accessibili all'osservazione dei telescopi ottici. (Le lunghezze d'onda ottiche hanno un valore caratteristico pari a poco meno di un micrometro, ossia un milionesimo di metro.) Di solito la radiazione nella banda millimetrica e in quella submillimetrica attraversa però le nubi come se fossero pressoché trasparenti e può di conseguenza fornire indizi sulle regioni dove nascono le stelle.

In secondo luogo, la radiazione nella banda submillimetrica può dare informazioni anche sulla natura del gas interstellare. Nonostante che la maggior parte di questo gas denso sia composta da molecole di idrogeno, quasi la metà di esso potrebbe essere costituita da una grande varietà di molecole in traccia. Di quali molecole si tratta? Quali sono le loro abbondanze relative? Le leggi della meccanica quantistica mettono a disposizione un sistema per rispondere a queste domande. In particolare, una data molecola può ruotare solo a un particolare insieme di frequenze discrete specifiche, le quali dipendono dalle proprietà della molecola.

Le molecole pesanti, per esempio, ruotano a frequenze relativamente basse e quelle leggere a frequenze più elevate. Quando una molecola compie una transizione da una velocità di rotazione permessa a un'altra, emette radiazioni di una frequenza caratteristica grazie alle quali è possibile identificare la particolare specie di molecola. Buona parte della radiazione delle molecole che ruotano nel gas interstellare si trova nella banda submillimetrica. Il rilevatore a effetto tunnel a semiconduttori è di conseguenza uno strumento ideale per la determinazione della composizione del gas.

Il «cuore» di un rilevatore a effetto tunnel a superconduttori è costituito da un sottile strato di materiale isolante (ossia una sostanza che non conduce elettricità) inserito a «sandwich» fra due superconduttori. Un superconduttore è un metallo, come piombo o niobio, che presenta proprietà notevolmente interessanti quando viene raffreddato a temperature vicine allo zero assoluto (zero kelvin, cioè -273 gradi Celsius). La proprietà forse più interessante in assoluto è che un superconduttore lascia passare l'elettricità senza alcuna resistenza. Un simile sandwich costituito da un superconduttore, un isolante e un superconduttore viene definito «giunzione a effetto tunnel a superconduttori».

Orbene, secondo le leggi della fisica classica il materiale isolante dovrebbe impedire alla corrente elettrica di passare attraverso la giunzione. La meccanica quantistica, però, offre una scappatoia: secondo questa importante teoria, un elettrone è rappresentato da un'onda di probabilità, e perciò la sua posizione è indeterminata. Di conseguenza esiste una piccola probabilità che un elettrone di un superconduttore possa comparire nell'altro superconduttore, come se fosse passato attraverso l'isolante in un tunnel.

Nel funzionamento di un rilevatore a effetto tunnel a superconduttori, le onde radio vengono fatte convergere su una giunzione a effetto tunnel a superconduttori. Nei superconduttori gli elettroni assorbono l'energia della radiazione; quelli di energia più elevata hanno una probabilità molto maggiore di attraversare l'isolante per effetto tunnel. Di conseguenza, la quantità di corrente elettrica che attraversa l'isolante aumenta con la quantità e il tipo delle onde radio che incidono sulla giunzione. Misurando la corrente si può risalire alla natura della radiazione elettromagnetica che stimola il dispositivo.

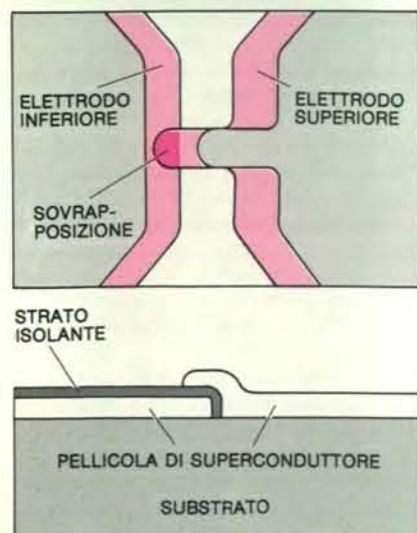
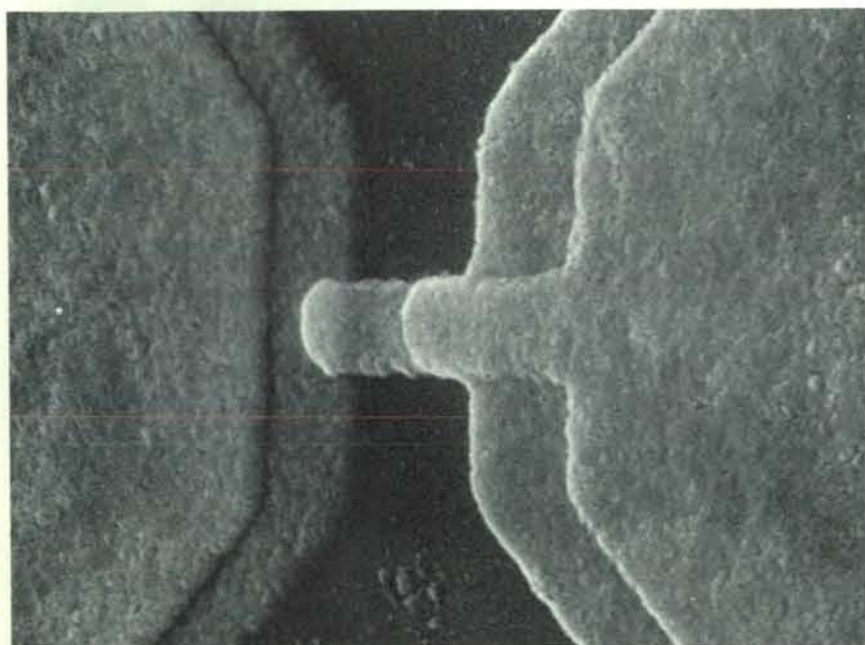
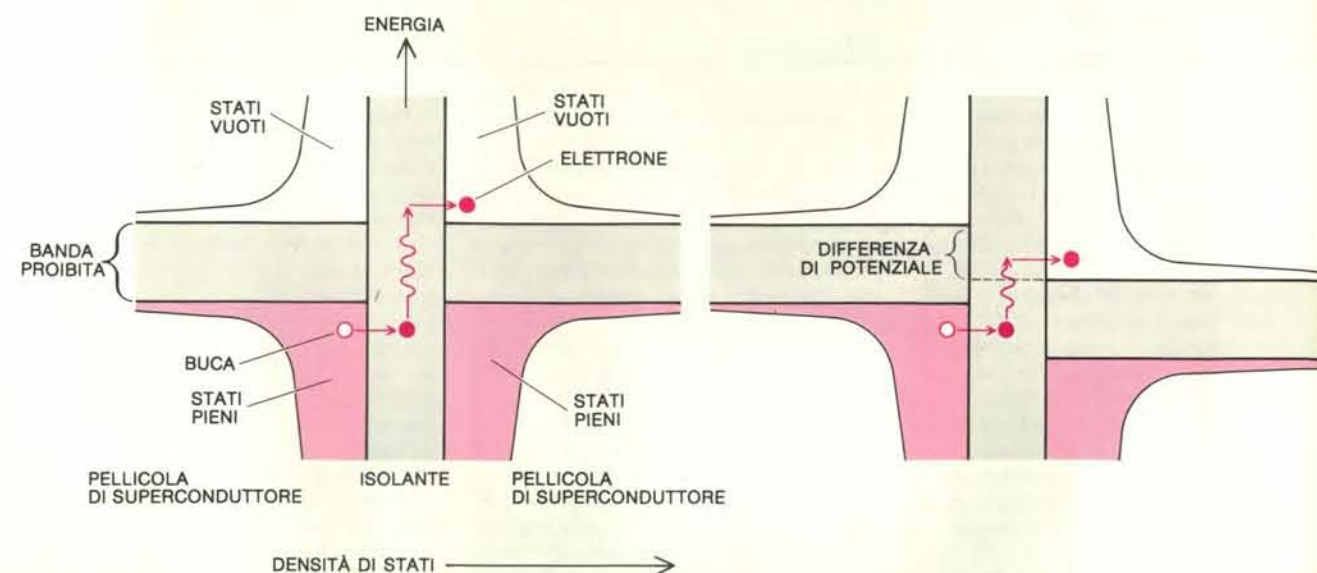
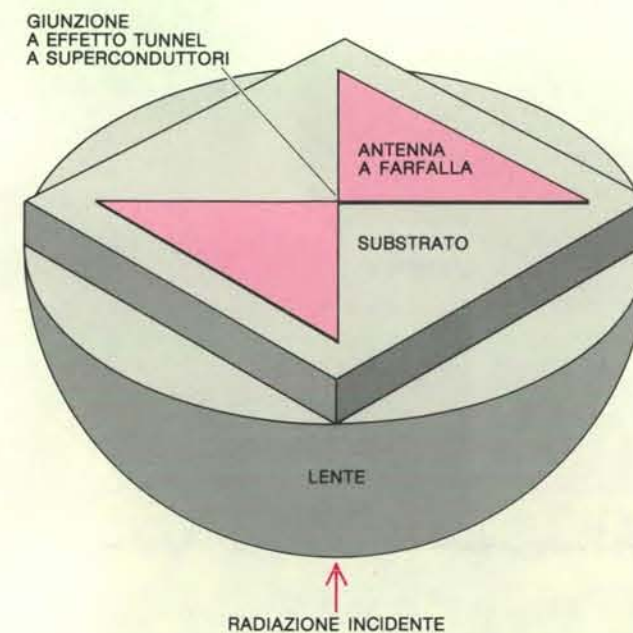
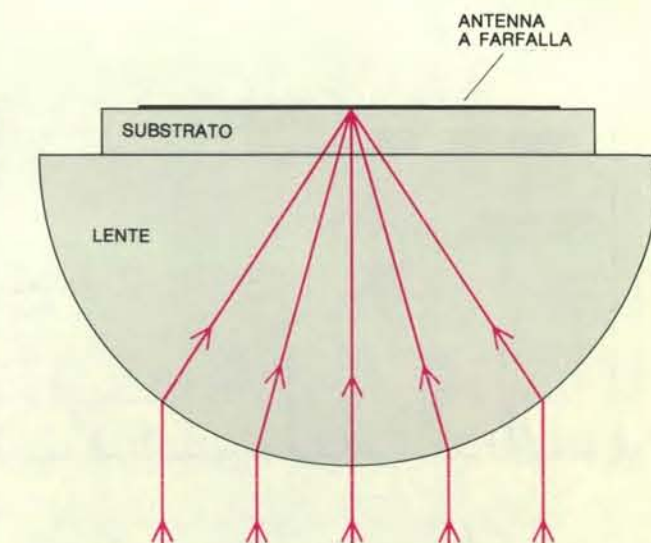
Nonostante le proprietà piuttosto fuori dall'ordinario del rilevatore a effetto tunnel a superconduttori, il principio fondamentale del suo funzionamento - la conduzione elettrica indotta dalla radiazione incidente - è lo stesso che sta alla base dei rilevatori convenzionali utilizzati negli os-

servatori per misurare la radiazione visibile. Per esempio, se immaginiamo una versione estremamente semplificata di un dispositivo ottico, vediamo che la luce proveniente da un telescopio viene concentrata su un pezzetto di materiale semiconduttore come il silicio. Gli atomi di un semiconduttore trattengono in maniera molto stretta quasi tutti i loro elettroni, con l'eccezione importante di quei pochi che vengono chiamati elettroni di valenza, o elettroni esterni. Sebbene gli elettroni di valenza non siano così strettamente legati come gli altri, di solito possono muoversi solamente da un atomo a un altro adiacente, partecipando alla formazione di legami chimici. La radiazione luminosa assorbita dal semiconduttore può tuttavia portare gli elettroni di valenza a un livello di energia più elevato, conosciuto come banda di conduzione: uno stato nel quale gli elettroni si possono muovere liberamente per tutto il materiale. Se si applica una differenza di potenziale elettrico ai capi del semiconduttore, gli elettroni fluiranno perciò da un capo all'altro del materiale. L'entità del flusso, ossia la corrente elettrica, è proporzionale alla quantità di luce incidente sul dispositivo: misurando la corrente si possono quindi dedurre la quantità e la natura della luce che è entrata nell'apertura del telescopio.

Un dispositivo di tipo convenzionale non è in grado però di rilevare onde radio. Il motivo va ricercato nel fatto che questa radiazione non possiede un'energia sufficiente a far salire gli elettroni di valenza dei semiconduttori nelle bande di conduzione. L'energia di qualsiasi onda elettromagnetica, infatti, è inversamente proporzionale alla sua lunghezza d'onda. Poiché un'onda radio millimetrica ha un valore di lunghezza d'onda pari a circa 1000 volte quello di un'onda luminosa, la sua energia è circa 1000 volte inferiore. L'energia richiesta per generare una corrente elettrica in un semiconduttore è pari a circa un elettronvolt, mentre quella posseduta da un'onda radio con una lunghezza d'onda di un millimetro è circa un millesimo di elettronvolt. Dal momento che un superconduttore è un metallo, esso possiede già molti elettroni nella banda di conduzione e non può essere utilizzato per rilevare la luce. I suoi elettroni, però, possono occupare stati energetici distinti. La quantità di energia necessaria per spostare gli elettroni da uno stato all'altro corrisponde all'energia di un'onda radio con una lunghezza d'onda di un millimetro, e questo spiega perché il rilevatore a effetto tunnel a superconduttori sia particolarmente adatto per applicazioni in campo radioastronomico.

L'esistenza di questi stati energetici distinti deriva dalla teoria della superconduttività, i cui elementi fondamentali furono esposti per la prima volta in un saggio del 1957 da John Bardeen, Leon N. Cooper e J. Robert Schrieffer, che allora lavoravano all'Università dell'Illinois a Urbana-Champaign. Secondo la loro teoria, in un superconduttore gli elettroni si legano formando le coppie di Cooper cioè, in altre parole, subiscono una forza di attrazione. Sebbene gli elettroni abbiano la stessa carica (negativa) e nello spazio libero si respingano reciprocamente, in un metallo essi possono dare l'impressione di attirarsi reciprocamente a causa dell'interazione con gli ioni dotati di carica positiva che vi si trovano. Gli ioni positivi, infatti, sono disposti in un reticolo che viene distorto ogni volta che un elettrone lo attraversa. Le distorsioni a loro volta attraggono un altro elettrone cosicché i due elettroni formano una coppia. Questa interazione di cooperazione porta alla comparsa di una banda di energie proibite tra gli stati permessi agli elettroni: al di sotto di questa banda tutti gli stati sono pieni, al di sopra sono vuoti.

In assenza di radiazione, quindi, gli elettroni di un rilevatore a effetto tunnel a superconduttori sono legati in coppie di Cooper. L'applicazione di onde radio alla giunzione scinde alcune delle coppie in uno dei due superconduttori e porta un elettrone di ciascuna coppia a un livello superiore di energia. Questi elettroni di energia superiore possono allora attraversare per effetto tunnel l'isolante e accedere agli stati vuoti dall'altra parte, aumentando così il flusso di corrente attraverso il dispositivo. Se il rilevatore è montato alla base di un telescopio, la corrente conterrà informazioni sull'oggetto astronomico sul quale il telescopio è puntato.

Le vicende che hanno portato dalla prima relazione di Bardeen, Cooper e Schrieffer alla realizzazione del rilevatore a effetto tunnel a superconduttori hanno visto la partecipazione di molti ricercatori. Buona parte del lavoro è stata compiuta per ragioni diverse dallo sviluppo di nuovi rilevatori. Verso l'inizio degli anni sessanta Ivan Giaver della General Electric Company dimostrò molte delle proprietà delle giunzioni a effetto tunnel a superconduttori. Dieci anni più tardi, gli studi compirono notevoli progressi quando ricercatori sia dell'IBM, sia degli AT&T Bell Laboratories cercarono di ricavare elementi per calcolatori a commutazione rapida dalle giunzioni a effetto tunnel a superconduttori. Si constatò che un dispositivo ad alta velocità deve essere straordinariamente piccolo, dell'ordine di un micrometro quadrato. Per costruire un elemento di questo tipo i gruppi dell'IBM e dei Bell Laboratories hanno dovuto in primo luogo mettere a punto una tecnica di microlitografia necessaria per poter incidere il dispositivo su un chip di silicio. La realizzazione di questo procedimento indusse altri ricercatori a occuparsi dello sviluppo di rilevatori per onde radio, dal momento che anche questi dispositivi devono essere eccezionalmente piccoli per poter registrare le frequenze relativamente elevate che sono caratteristiche della radiazione millime-

Il «cuore» del rilevatore a effetto tunnel a superconduttori è costituito da un sottile strato di materiale isolante inserito tra due superconduttori. Il dispositivo, visibile nella fotografia al microscopio elettronico a scansione, è costituito da due sottili pellicole in una lega di piombo; il piombo diventa superconduttore a temperature prossime allo zero assoluto (−273 gradi Celsius). La regione di sovrapposizione delle pellicole (*si vedano a sinistra la vista in pianta e la sezione*) costituisce il rilevatore, che ha un'area di circa un micrometro (un milionesimo di metro) quadrato. Lo strato isolante tra le due pellicole di piombo è stato realizzato esponendo la superficie superiore della pellicola inferiore all'aria e facendovi sviluppare sopra uno strato naturale di ossido. Il dispositivo è stato realizzato da Ronald E. Miller degli AT&T Bell Laboratories. La tecnica di preparazione, nota come evaporazione a doppio angolo, fa sembrare entrambe le pellicole doppie. Tuttavia, durante il funzionamento del rilevatore, ha importanza solo la regione di sovrapposizione.

L'effetto tunnel degli elettroni è alla base del funzionamento del rilevatore a effetto tunnel a superconduttori. Un'interazione di cooperazione tra gli elettroni di un superconduttore porta alla formazione di una banda proibita fra gli stati di energia permessi agli elettroni (*a sinistra*). Sotto questa banda tutti gli stati sono pieni; sopra sono vuoti. L'applicazione di onde radio al rilevatore fa salire alcuni elettroni di uno dei superconduttori a un'energia superiore. Questi elettroni possono attraversare l'isolante per effetto tunnel e accedere agli stati vuoti dall'altra parte. Misurando il flusso di elettroni, ossia la corrente elettrica, attraverso l'isolante si possono dedurre le proprietà della radiazione incidente sul dispositivo e così risalire alla natura della sorgente della radiazione. Si può sintonizzare il rilevatore su una lunghezza d'onda particolare della radiazione applicando una differenza di potenziale attraverso la giunzione (*a destra*) ed elevando così i livelli energetici di un lato, rispetto a quelli dell'altro.

L'antenna a farfalla «alimenta» il rilevatore a effetto tunnel a superconduttori che si trova al suo centro (*a sinistra*) con le onde radio provenienti dal telescopio. Un'antenna tipica è lunga pochi millimetri e spessa un decimo di micrometro. Il metallo viene fatto depositare per evaporazione sul substrato, di solito quarzo o silicio, e la matrice si ottiene con gli stessi procedimenti fotolitografici che si usano per i circuiti integrati. Le antenne di questo tipo sono facili da realizzare, ma non scevre da problemi. In particolare interagiscono fortemente con i modi del substrato, ossia con le onde che vi si propagano. L'interazione può portare alla perdita anche dell'80 per cento della radiazione incidente. Una soluzione ideata dagli autori consiste nell'attaccare una lente dello stesso materiale del substrato alla sua faccia posteriore (*a destra*). La lente focalizza la radiazione sull'antenna a farfalla eliminando i modi del substrato. Un dispositivo del genere può funzionare su una vasta gamma di frequenze.

trica e della radiazione submillimetrica.

Le prime relazioni su rilevatori funzionanti a effetto tunnel a superconduttori con rumore, cioè disturbi indesiderati, sorprendentemente basso furono presentate contemporaneamente nel 1979 da Gerard J. Dolan, uno di noi (Phillips) e David P. Woody dei Bell Laboratories, da Paul L. Richards e Tek-Ming Shen dell'Università della California a Berkeley e da Richard E. Harris e Frances L. Lloyd del National Bureau of Standards. Sempre nel 1979 anche Staffan Rudner e Tord Claeson dell'Istituto di tecnologia Chalmers in Svezia stesero una relazione su un dispositivo simile. La storia, comunque, non è tutta qui. Altri ricercatori hanno fornito contributi importanti allo sviluppo dei rilevatori a superconduttori e tra questi vi sono John R. Tucker dell'Università dell'Illinois, Aly H. Dayem e Raymond J. Martin dei Bell Laboratories.

Attualmente in molti laboratori di tutto il mondo si ricorre all'impiego di differenti tecniche litografiche, ossia procedimenti di stampa, per realizzare i rilevatori che sono fabbricati normalmente con leghe di piombo. Lo strato isolante tra i due pezzi di piombo ha uno spessore di circa due nanometri (miliardesimi di metro). Questo spessore estremamente sottile viene ottenuto esponendo semplicemente la superficie superiore del pezzo inferiore di metallo all'aria e facendo sì che vi si sviluppi sopra uno strato naturale di ossido. Il piombo diventa superconduttore a circa otto kelvin, e così gli strumenti vengono fatti funzionare in un criostato che contiene elio liquido. (L'elio diventa liquido a 4,2 kelvin.) I progressi futuri comporteranno senza dubbio l'impiego di superconduttori al niobio o al nitruro di niobio, che sono più difficili da produrre, ma dovrebbero essere meno suscettibili alla degradazione chimica e possedere una sensibilità maggiore.

Non abbiamo ancora affrontato il problema di come sia possibile «incanalare» onde radio provenienti da un telescopio enorme, con un diametro di dieci metri, in un minuscolo rilevatore a effetto tunnel a superconduttori che misura un micrometro di larghezza. Mentre è comprensibile che il rilevatore debba essere piccolo per rispondere alle alte frequenze radio, è paradossale che le sue dimensioni siano tanto più piccole di una lunghezza d'onda che non è possibile focalizzarvi direttamente l'onda come si fa nel caso dei rilevatori presenti nei telescopi ottici e in quelli infrarossi.

La soluzione sta nel costruire una piccola antenna di alimentazione che catturi le onde focalizzate dal telescopio e le porti alla giunzione a effetto tunnel a superconduttori. Per le lunghezze d'onda millimetriche l'antenna è di solito una tromba di metallo collegata a un tubo cavo di metallo. La giunzione si trova su un piccolo pezzetto di silicio montato all'interno del tubo. Sebbene le guide d'onda di metallo cave trasferiscano l'energia al rilevatore in modo efficiente, hanno pur sempre delle limitazioni. La guida d'onda è alta solamente un quarto di una lunghezza d'onda e questo rende difficile costruire le guide e montare i rilevatori per lunghezze d'onda submillimetriche.

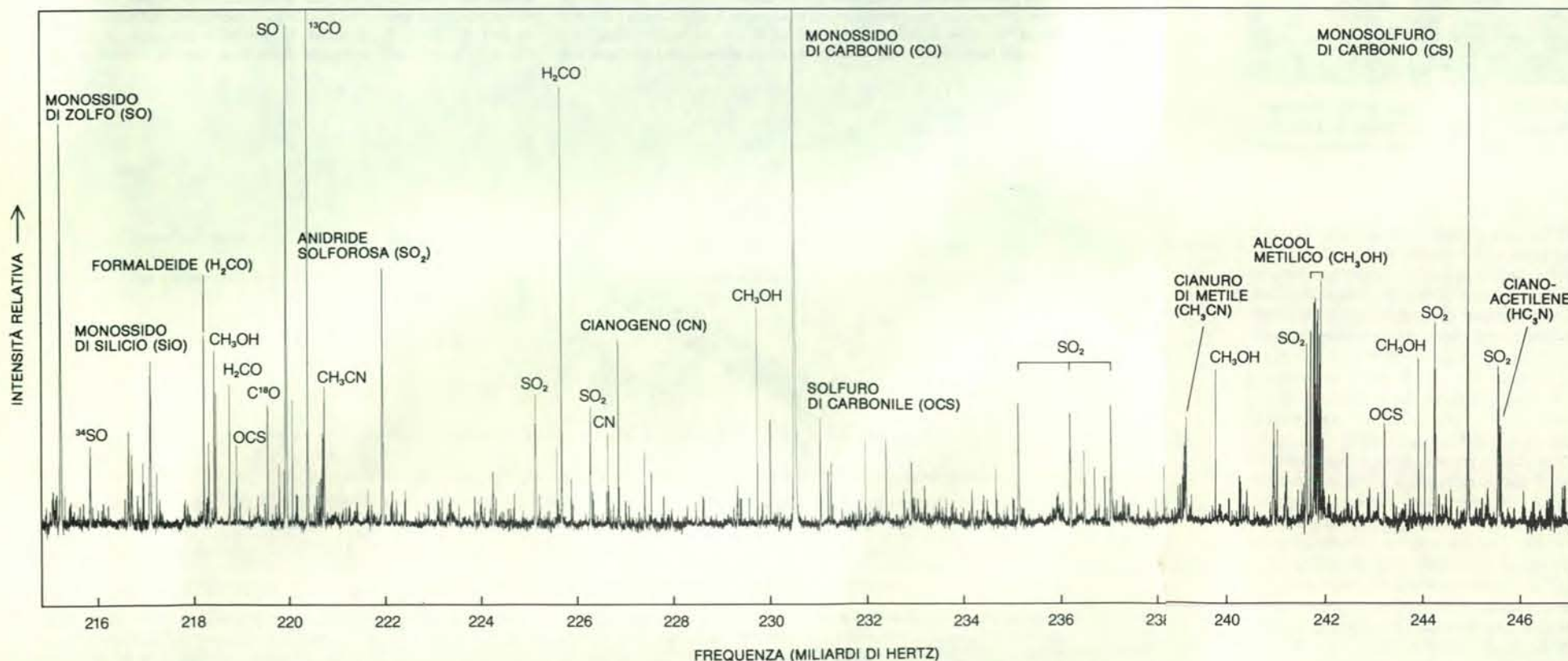
Per superare all'origine questo problema abbiamo fatto sì che l'antenna di alimentazione sia parte di un circuito integrato che contiene la giunzione a effetto tunnel a superconduttori. Invece di un tubo di metallo lavorato, l'antenna di alimentazione è una pellicola di metallo ottenuta per evaporazione e il cui spessore è un decimo di micrometro. Le matrici sono realizzate con gli stessi procedimenti fotolitografici che vengono usati nella fabbricazione dei circuiti integrati, ed è facile quindi riuscire a tenere sotto controllo dimensioni anche di un micrometro. Inoltre vi è un vantaggio in più: è possibile fabbricare contemporaneamente sullo stesso substrato un gran numero di antenne e di rilevatori, ottenendo così un gruppo (o una schiera) di elementi per la formazione delle immagini analogo a quello di una telecamera. Una schiera di questo tipo farebbe accelerare notevolmente la cartografia delle grandi sorgenti astronomiche.

Altre difficoltà, però, affliggono le antenne a circuito integrato, le quali interagiscono intensamente con i modi del substrato, ossia con le onde che si propagano nel substrato. L'interazione può dar luogo a una perdita anche dell'80 per cento della radiazione incidente. Un metodo per affrontare questo problema consiste nell'eliminare i modi del substrato rendendo quest'ultimo molto sottile. Oggi è possibile produrre membrane di silicio e di nitruro di silicio senza substrato aventi un'area di un centimetro quadrato e uno spessore inferiore a un micrometro. Attualmente, siamo impegnati nella realizzazione di una schiera di antenne per lunghezze d'onda submillimetriche su una pellicola di nitruro di silicio.

Un altro metodo consiste nel collocare dietro il substrato una lente dello stesso materiale del substrato, la quale ne elimina i modi e sfrutta il fatto che le antenne sono molto più sensibili alla radiazione proveniente dal substrato che a quella proveniente dall'aria. Dean P. Neikirk dell'Università del Texas ad Austin adottò questo metodo diversi anni fa quando era laureando al California Institute of Technology (Caltech). Più recentemente Michael J. Wengler, uno dei nostri laureandi del California Institute of Technology, ha accoppiato una lente da substrato con un'antenna a farfalla (*bow-tie antenna*) contenente al centro una giunzione a effetto tunnel a superconduttori. Le prestazioni di questo dispositivo si sono dimostrate eccellenti su una gamma di frequenze molto più ampia di quella di qualsiasi altro ricevitore per radioastronomia.

Quando si fa funzionare un rilevatore a effetto tunnel a superconduttori in un telescopio, si «mescolano» le onde radio con un fascio di riferimento e si focalizza il fascio risultante sull'antenna del substrato. Questo procedimento è noto come rilevamento a eterodina e fa aumentare notevolmente il potere di risoluzione spettrale dello strumento. Uno strumento chiamato oscillatore locale genera il fascio di riferimento che è formato da onde radio di frequenza costante. Viene osservata e analizzata la nota di battimento risultante, ossia una frequenza intermedia, per mettere in luce le informazioni relative all'oggetto astronomico osservato.

I rilevatori a effetto tunnel a superconduttori hanno già offerto alcuni contributi importanti all'astronomia. Come ha messo in evidenza il telescopio per onde millimetriche del Caltech, che è dotato di un dispositivo di questo tipo, la nube gassosa intorno a Orione è composta da almeno 28 specie molecolari diverse. Le misurazioni sono state eseguite da uno dei nostri laureandi, Geoffrey A. Blake, insieme ai ricercatori Edmund C. Sutton e Colin R. Masson e a uno di noi (Phillips), e sono state identificate più di 500 righe spettrali emesse dalle molecole. Il rilevatore, che è stato costruito da Ronald E. Miller presso i Bell Laboratories, è tanto sensibile che permette agli astronomi di identificare le più rare tra le molecole, persino quelle che contengono deuterio, l'isotopo pesante dell'idrogeno. La capacità di misurare l'abbondanza di tante specie molecolari avrà senza dubbio un'influenza notevole nell'importante campo della chimica interstellare.

I rilevatori a effetto tunnel a superconduttori possono determinare anche il moto relativo delle molecole in avvicinamento o in allontanamento dall'osservatore rilevando lo spostamento di frequenza della radiazione emessa. (Questo spostamento è noto come «effetto Doppler», un fenomeno familiare a chiunque abbia notato che mentre passa un treno il tono del suo fischio cambia da alto a basso.) In questo modo si possono distinguere l'una dall'altra singole nubi di gas all'interno di una località di formazione di nuove stelle. Nella nube di Orione, per esempio, le molecole di metanolo hanno righe spettrali molto strette, mentre le molecole di anidride solforosa hanno righe larghe. La ragione di questa differenza sta nel fatto che il metanolo si trova principalmente nelle regioni quiescenti della nube, mentre l'anidride solforosa si trova soltanto nelle regioni attive dal punto di vista fisico e chimico. I processi di ricca formazione di stelle han-

Una nube gassosa di molecole circonda parte della costellazione di Orione (*a destra*). Anche se per la massima parte è composto da idrogeno atomico, il gas contiene anche una grande varietà di molecole in traccia. Gli astronomi le identificano misurando le frequenze della loro emissione radiativa caratteristica. Lo spettro molecolare della nube di Orione è raffigurato qui sopra. Tutti i dati sono stati ottenuti con un telescopio per onde millimetriche dotato di un rilevatore a effetto tunnel a superconduttori. Il gas contiene almeno 28 specie molecolari diverse (non indicate).

L'interferometro di Owens Valley, che è gestito dal Caltech, è composto da tre telescopi per onde millimetriche collegati elettronicamente e su ognuno è montato un rilevatore a effetto tunnel a superconduttori. L'immagine in basso a pagina 70 è stata realizzata con questo interferometro. I telescopi sono stati costruiti sotto la direzione di Robert B. Leighton e nell'insieme hanno un diametro efficace di parecchie centinaia di metri.

no luogo in queste regioni attive, dando origine a forti venti e flussi di materia verso l'esterno.

Nonostante la sensibilità dei rilevatori, tuttavia, molti oggetti astronomici sono troppo piccoli per poter essere osservati con un telescopio convenzionale. A causa della diffrazione della radiazione, infatti, la migliore risoluzione angolare di un telescopio di 10 metri di diametro per radiazione di tre millimetri è circa un minuto d'arco, mentre le regioni di formazione di stelle, i nuclei galattici e altri oggetti lontani possono avere strutture sulla scala dei secondi d'arco. Per aggirare questa difficoltà i radioastronomi collegano elettronicamente diversi telescopi piccoli, creando in pratica un telescopio molto più grande. È il metodo dell'interferometria. L'interferometro per onde millimetriche di Owens Valley del California Institute of Technology sfrutta proprio questa tecnica. L'apparecchiatura consiste di un gruppo di tre telescopi per onde millimetriche, costruiti da Robert B. Leighton, i quali hanno un diametro di 10,4 metri ma che, collegati insieme, hanno un diametro efficace di diverse centinaia di metri. Ciascun telescopio è equipaggiato con un rilevatore a effetto tunnel a superconduttori alla base.

Probabilmente gli oggetti più entusiasmanti osservati finora nel cielo con la schiera interferometrica sono le condensazioni di gas nei nuclei delle galassie vicine. Concentrando l'attenzione sull'emissione estremamente intensa proveniente dalla prolifica molecola di monossido di carbonio, abbiamo realizzato una carta dell'emissione del gas nel nucleo della galassia M82. Le curve di livello del monossido di carbonio delineano una regione al centro della galassia, dove si osservano molti resti di supernova e dove si stanno formando proprio adesso stelle di grande massa. Le immagini riprese nel visibile rivelano che il gas si sta allontanando da questa regione di intensa attività.

Sono in costruzione anche telescopi per la banda submillimetrica. Per evitare gli effetti di attenuazione dovuti alle molecole di acqua e di ossigeno dell'atmosfera terrestre, queste apparecchiature vengono installate ad alta quota. Il gruppo del California Institute of Technology e altri stanno costruendo telescopi per onde submillimetriche in località montane come il Mauna Kea, ad Hawaii. Sono state effettuate misurazioni anche dal Kuiper Airborne Observatory della National Aeronautics and Space Administration (NASA), che è costituito da un telescopio ottico di un metro di diametro portato a una quota di circa 12 000 metri da un aereo da trasporto.

Nel lungo termine gli astronomi sperano che l'intero spettro submillimetrico, comprese le regioni oscurate dall'acqua e dall'ossigeno presenti nell'atmosfera terrestre, divenga accessibile mediante un grande riflettore installabile nello spazio (LDR, da *large deployable reflector*), il quale è ora in programmazione alla NASA. Tutti questi nuovi progetti di telescopi comprendono, o prevedono di comprendere, rilevatori a effetto tunnel a superconduttori.

Caccia e raccolta nelle foreste dell'Europa postglaciale

Spesso considerate poco più di un'introduzione all'agricoltura primitiva, sono state di recente rivalutate e si ritiene oggi che abbiano rappresentato uno sviluppo parallelo all'agricoltura e talvolta altrettanto produttivo

di Marek Zvelebil

Circa 10 000 anni fa i ghiacci che per molti millenni avevano ricoperto una buona parte dell'Europa e dell'Asia cominciarono rapidamente a fondersi. Al ritirarsi dei ghiacciai, nuovi ambienti si formarono nell'Eurasia settentrionale. Steppa e tundra, che erano state comuni durante le grandi glaciazioni del Pleistocene, furono colonizzate da foreste di conifere e caducifoglie. Nelle regioni ora ricoperte da foreste si affermò un nuovo modo di vita. Durante il Pleistocene gli abitanti dell'Eurasia settentrionale erano vissuti in gran parte cibandosi di selvaggina di grossa taglia, alla quale davano la caccia attraverso la steppa ghiacciata. La foresta che sostituì la steppa era un ecosistema complesso contenente un'associazione molto più ricca di animali e di piante commestibili, e le bande di cacciatori-raccoglitori del periodo postglaciale sfruttarono appieno la nuova abbondanza, raccogliendo cibi vegetali e cacciando, oltre ai mammiferi di grossa taglia, mammiferi di piccola taglia e uccelli acquatici.

Il modo di vita del periodo postglaciale, fondato sulla caccia e sulla raccolta, durò finché non fu sostituito dall'agricoltura, trasformazione che nell'Europa settentrionale avvenne circa 5000 anni fa. Il periodo di 5000 anni fra il ritiro dei ghiacciai e l'introduzione dell'agricoltura è stato considerato tradizionalmente un modesto interludio, nella convinzione che la caccia e raccolta fosse un modo di sussistenza rozzo, capace di fornire i mezzi di sostentamento solo a piccoli gruppi sempre in movimento per procurarsi faticosamente di che vivere nella foresta. Nella concezione tradizionale, la caccia e raccolta non era certo un modo di vita in grado di reggere la concorrenza con l'agricoltura, una volta che questa fu introdotta. Era perciò opinione comune che l'agricoltura si fosse diffusa rapidamente dal Vicino Oriente alle regioni d'Europa ricoperte da foreste, dove fu introdotta da agricoltori colonizzatori o adottata da bande di cacciatori-raccoglitori che ne percepirono rapidamente l'importanza per migliorare le loro misere condizioni di vita.

Quest'opinione ha cominciato a modificarsi di recente, con la progressiva messa a fuoco della complessità e della produttività delle bande di cacciatori-raccoglitori postglaciali. Ed è stato chiarito che, in ambienti favorevoli, la caccia e raccolta poteva fornire i mezzi di sussistenza a popolazioni molto più dense di quanto si pensasse in precedenza. In alcune aree le bande di cacciatori-raccoglitori conducevano una vita relativamente sedentaria in insediamenti permanenti, contraddistinti da una cultura sviluppata e da elementi di differenziazione sociale. Queste comunità non adottarono prontamente l'agricoltura e fu chiaro che, in certe zone dell'Europa settentrionale, l'avvento dell'agricoltura subì un notevole ritardo e le bande di cacciatori-raccoglitori adottarono il nuovo sistema di vita solo quando furono costrette a farlo da crisi conseguenti a mutamenti del clima. La tenace persistenza della caccia e raccolta per molto tempo dopo che essa «sarebbe dovuta» sparire è uno degli elementi che stanno contribuendo a una rivalutazione radicale della caccia e raccolta postglaciale. Penso che, quando questa rivalutazione sarà completa, la caccia e raccolta nelle foreste postglaciali sarà considerata uno sviluppo parallelo all'agricoltura e, almeno per qualche tempo, altrettanto vitale come mezzo di sussistenza.

La zona a foreste dell'Eurasia, a cui si applicano le nuove conclusioni, non è ecologicamente omogenea. A settentrione le foreste sono formate principalmente da conifere. A latitudini più temperate prevalgono specie di caducifoglie, o un misto di specie dell'uno e dell'altro tipo. La distribuzione di queste componenti è variata in misura considerevole durante le fluttuazioni climatiche degli ultimi 10 000 anni. Al tempo del massimo sviluppo delle società postglaciali di cacciatori-raccoglitori, la zona ricoperta da foreste era estesa più a nord di quanto non sia oggi. Alcune aree che oggi sono semidesertiche o ricoperte da praterie erano coperte allora da steppe miste a foreste e da boschi.

Per tener conto di queste variazioni, adotterò per la zona a foreste una definizione geografica lata anziché quella ecologica rigorosa. Ai miei fini, tale zona copre la maggior parte dell'Europa a nord della latitudine di 40 gradi. (Il quarantesimo parallelo passa per la Spagna centrale, l'Italia meridionale, la penisola greca e la Turchia settentrionale, prima di raggiungere il Mar Caspio.) Uno dei caratteri più significativi della regione così definita è che, durante il post-Pleistocene, molte delle sue risorse erano concentrate in prossimità di bacini d'acqua: fiumi, laghi e mari.

Un altro aspetto importante dell'ecologia della zona a foreste è stato il fatto che, nel corso del tempo, la quantità di risorse disponibili subì vistose fluttuazioni. Le fluttuazioni furono di vari tipi. Quelle a lungo termine dipendevano dalla struttura ecologica della stessa foresta eurasiatica. Quivi la catena alimentare era dominata da specie come la renna, la lepre e selvaggina alata, che si nutrivano di piante e fornivano a loro volta cibo a predatori come orsi e linci. Molte di queste specie erano opportuniste (in senso zoologico). Riproducendosi in breve tempo e avendo prole abbondante, tendono a crescere rapidamente come popolazione in concomitanza di una maggiore disponibilità di cibo. Se le prede aumentano, aumentano anche i predatori. Spesso il risultato finale è una drastica diminuzione delle prede. La presenza di molte specie opportuniste tra queste ha indubbiamente prodotto variazioni imprevedibili ed estreme nelle risorse di cibo dei vari gruppi di cacciatori-raccoglitori.

Il secondo tipo di fluttuazione, più prevedibile e più a breve termine, è la variazione che si accompagna al ciclo delle stagioni. Le foreste nella zona temperata settentrionale sono luoghi i cui caratteri variano spiccatamente con le stagioni. Il tipo e la quantità di cibo disponibili mutano in modo brusco da una stagione all'altra. Gran parte della variazione è dovuta al fatto che alcune tra le fonti di cibo più significative sono costituite da specie migratorie, fra le quali molte specie di uccelli acquatici e pesci anadromi come il salmone, che tornano ogni anno dal mare a riprodursi nei corsi d'acqua in cui sono nati. La periodicità di queste migrazioni comporta che alcune specie siano disponibili solo per brevi periodi dell'anno, generalmente in primavera e in autunno. In altri periodi, specialmente in quello più freddo dell'inverno, il cibo può scarseggiare.

In che modo gli abitanti dell'Europa postglaciale poterono far fronte a questa variabilità? Una possibile risposta a questa domanda potrebbe essere stata data dall'agricoltura. Questa non fu però, in origine, un'opzione per gli abitanti della zona a foreste. La ragione è che l'agricoltura, quale si sviluppò nel Vicino Oriente, si fonda sia sulla coltivazione dei campi, sia sull'allevamento del bestiame. Le due attività dipendono a loro volta dalla presenza di una combinazione di specie domesticabili: animali da condurre al pascolo e piante da coltivare. Nel Vicino Oriente i cereali fornirono la base per l'agricoltura, mentre pecore e capre ebbero la stessa funzione per la pastorizia. La zona a foreste settentrionale non possedeva una combinazione così propizia. Alcune specie potevano essere addomesticate, ma tendevano a presentarsi isolatamente e non come parte della varietà che era necessaria per l'insediamento dell'agricoltura.

Ma l'agricoltura è solo una delle possibili strategie per aumentare la produzione di risorse e per rimediare all'instabilità intrinseca dell'ambiente. Sono possibili altre strategie, e quel che esse condividono con l'agricoltura è un maggiore investimento di forza-lavoro per ottenere cibo. Questo maggior investimento potrebbe cominciare, per esempio, con l'accumulo del maggior numero di informazioni possibile sul ciclo vitale di una specie. Fon-

Proveniente da un sito in prossimità del villaggio di Niskala, nella Finlandia settentrionale, questo manufatto di ardesia finemente striato, con la testa di un alce, fu intagliato nel III millennio a.C. da un membro di una banda di cacciatori-raccoglitori della tarda Età della pietra. Alcuni archeologi pensano che fosse un coltello, ma l'autore ritiene che avesse una funzione principalmente sociale: il suo possesso doveva essere un mezzo per elevare il proprio rango. Nel periodo postglaciale avanzato oggetti simbolo di rango come questo apparvero in gran numero nelle foreste europee, il che potrebbe indicare un accentuarsi della competizione sociale all'interno dei gruppi locali di cacciatori-raccoglitori. Quest'oggetto fa parte di una collezione del Museo nazionale di Helsinki.

La zona a foreste eurasiatica, nella quale sorsero culture fondate su tecniche complesse di caccia e raccolta, comprende varie componenti. Gran parte dell'Europa settentrionale e dell'Asia settentrionale è coperta da foreste di conifere, da foreste di caducifoglie o da foreste miste, costituite dai due tipi di albero (*in colore*). La parte più a nord di quest'area è, invece, ricoperta da foreste composte esclusivamente da conifere (*area tratteggiata*). Nel corso degli ultimi 10 000 anni l'estensione della foresta di conifere e quella della zona a foreste complessiva hanno presentato fluttuazioni considerevoli. Anche per questa ragione l'autore ha deciso di usare una definizione lata della zona a foreste, includendovi la maggior parte dell'Europa a nord del 40° parallelo.

La distribuzione delle risorse nella zona a foreste è diseguale: animali e piante che forniscono cibo sono concentrati presso bacini d'acqua. Nell'Europa settentrionale e occidentale le aree più ricche si trovano vicino all'Atlantico (*in colore*). All'interno di questa regione le zone più produttive sono lungo certe linee di costa (*in colore più intenso*). Nell'Europa orientale le risorse sono concentrate nei bacini del Volga, del Dnepr e del Dnestr, oltre che nella pianura padana e nella valle del Danubio (*in grigio*). Sulle ricche risorse di queste aree si basarono culture fondate su tecniche complesse di caccia e raccolta, che si specializzarono nello sfruttamento di specie acquatiche.

dandosi su tale conoscenza si potrebbero sviluppare metodi speciali per catturare la preda. È così possibile sfruttare una specie come fonte di cibo in modo che le sue rese siano maggiori pur senza procedere a una completa domesticazione.

La documentazione archeologica suggerisce che questa forma di specializzazione sia stata la strategia adottata da molti gruppi postglaciali. Gli organismi nel cui sfruttamento questi gruppi si specializzarono erano generalmente acquatici perché la tendenza verso la specializzazione si delineò nel contesto di uno spostamento totale dalle risorse della terraferma a quelle dell'acqua. All'inizio del periodo postglaciale le principali risorse di cibo erano costituite, nella maggior parte dei siti, da ungulati, come cervi o bovini selvatici, e da altra selvaggina di grossa taglia. In seguito, nel corso dell'epoca postglaciale, vi fu una tendenza all'insediamento di siti specializzati dove una frazione significativa della dieta era costituita da una o poche specie. In generale queste erano acquatiche: pesci, crostacei, uccelli acquatici migratori o mammiferi marini come le foche. Il passaggio a risorse acquatiche non ebbe luogo simultaneamente in tutte le foreste eurasiatiche. Negli Urali meridionali i suoi inizi possono essere scoperti fra resti che risalgono al periodo compreso fra il 7000 e il 5000 a.C.; nelle gole del Danubio, in Iugoslavia, la data corrispondente è attorno al 5500 a.C.; nella penisola scandinava al 4000 a.C. e nella regione del Baltico orientale al 3000 a.C.

La cultura della «ceramica ad alveoli» (vasi con pozzetti, o alveoli, e con impressioni a pettine), nella Svezia meridionale, è un buon esempio di economia specializzata. I gruppi umani che vi appartenevano, e che fiorirono lungo le coste svedesi nel III millennio a.C., vissero cibandosi principalmente di foche dagli anelli. Vari adattamenti culturali significativi migliorarono l'efficienza con cui davano la caccia alle foche. Avevano, infatti, una tecnologia dell'arpone altamente sviluppata, che aumentava le probabilità di uccidere una foca se questa era a tiro. Dopo la caccia, olio e grasso di foca venivano riposti in grandi vasi di argilla, forse progettati specificamente per quello scopo. Persino la localizzazione delle comunità della ceramica ad alveoli rifletteva la specializzazione della caccia alla foca: le foche dagli anelli venivano cacciate in primavera sul ghiaccio costiero, dove procreano i piccoli, e gli insediamenti dei cacciatori venivano allestiti in prossimità di piccole baie e altre insenature costiere che, per la loro morfologia, davano facile accesso ai concentramenti di piccoli sui ghiacci.

La cultura della ceramica ad alveoli esemplifica vari caratteri essenziali della strategia specialistica. La massimizzazione del numero di prede durante il breve periodo in cui sono disponibili implica spesso la produzione di utensili specializzati, come l'arpone, la cui tecnologia è propria dei cacciatori di questa cultura. Gli utensili specializzati richiedono in genere un tempo più lungo per essere prodotti e riparati rispetto agli utensili di uso più generico, ma ciò non può avvenire durante la prima stagione di caccia. Per gli specialisti diventa quindi cruciale mettere in preventivo periodi di tempo da dedicare alla manutenzione degli utensili, periodi che dovranno essere trovati nelle stagioni morte, prima e dopo la caccia. Inoltre, se le prede uccise in un periodo di tempo relativamente breve serviranno a fornire gran parte della dieta per un periodo esteso, la loro conservazione diventerà cruciale. La cultura della ceramica ad alveoli era uno solo dei molti gruppi di cacciatori-raccoglitori postglaciali che possedevano metodi attendibili per la conservazione dei cibi.

La distribuzione del tempo fra le varie attività presso i cacciatori-raccoglitori della zona a foreste eurasiatica è del tutto diversa che ai tropici. In ciascuno di questi due istogrammi l'area dei rettangoli in grigio corrisponde al tempo impiegato nella preparazione degli utensili e l'area all'interno delle linee in colore corrisponde al tempo dedicato alla caccia e raccolta. In un ipotetico ambiente tropicale (*in alto*), la quantità di cibo disponibile non presenta grandi fluttuazioni, cosicché la preparazione degli utensili può essere distribuita sull'intera durata dell'anno. In un ipotetico ambiente forestale eurasiatico (*in basso*), la quantità di cibo è soggetta a variazioni considerevoli con le stagioni. La preparazione di utensili deve essere concentrata in periodi di inattività, in modo da dedicare il maggior tempo possibile alla caccia e raccolta quando vi è la massima disponibilità di cibo animale e vegetale. Questa situazione incoraggia la produzione di utensili specializzati per potenziare la cattura di prede durante il periodo più favorevole alla caccia.

La tecnologia specializzata, particolarmente in combinazione con metodi per la conservazione degli alimenti, contribuì a stabilizzare la disponibilità di cibo. Ma la strategia specialistica non era esente da rischi. Il fatto di concentrarsi su una fonte di cibo rese il gruppo vulnerabile a qualsiasi fluttuazione nella popolazione di quella specie. A causa dell'instabilità dell'ambiente forestale, era probabile che, a lungo termine, la popolazione di una qualsiasi specie presentasse delle oscillazioni. Una soluzione consisteva nel diversificare la dieta aumentando la varietà di specie vegetali e animali che ne facevano parte.

La diversificazione accentuò il contrasto fra i cacciatori-raccoglitori postglaciali e i loro predecessori del Paleolitico superiore, nel Pleistocene. Molti gruppi eurasiatici del Paleolitico superiore dipendevano da alcune specie di animali da caccia di grossa taglia, come renne, cavalli selvaggi e mammut lanosi. All'inizio del periodo postglaciale questa gamma ristretta aumentò in misura considerevole. La «cultura maglemosiana», così chiamata dalla torbiera danese di Maglemose in cui furono trovati i resti originari della cultura, ci offre un buon esempio di questa diversificazione. Durante l'VIII e il VII millennio a.C. essa si estese su una fascia comprendente la Gran Bretagna orientale, il bacino del Mare del Nord (che allora era prosciugato), la Danimarca, la Polonia settentrionale e la regione del Baltico orientale. La sua popolazione si nutriva di selvaggina di grossa taglia, come uri (bovini selvatici), cervi rossi, caprioli e cinghiali, pescava pesci, tra i quali lucci e salmoni, dava la caccia alle foche e agli uccelli migratori, oltre che alla selvaggina di piccola taglia come le lepri e i castori. La loro dieta comprendeva anche una quantità di prodotti vegetali, fra i quali nocciole, castagne d'acqua e, in alcune aree, persino ninfee.

Proprio come la strategia specialistica portò con sé utensili destinati a compiti sempre più ristretti, la strategia della diversificazione richiese utensili applicabili con uguale efficienza a molti compiti. Una tale tecnologia venne sviluppata rapidamente e la sua componente più importante fu un tipo di piccola lama in pietra chiamata microlito. I microliti, che si trovano in quasi tutti i siti postglaciali, hanno in generale una lunghezza compresa tra uno e due centimetri e una forma variabile da trapezoidale a triangolare e a lunata (a mezza luna). Essi venivano prodotti spezzando una lama più grande, che spesso era di selce. Dopo essere state staccate dalla lama più grande, le schegge venivano lavorate fino a far loro assumere la forma voluta. Una volta ricevuta la forma definitiva, venivano spesso innestate su un manico di legno, di corno o d'osso. Le grandi virtù della tecnologia dei microliti consistono nella versatilità delle microlame e nella facilità con cui potevano essere riparate. Una lama poteva essere tolta facilmente dall'impugnatura e affilata o sostituita da una nuova lama. In questo modo un gruppetto di cacciatori poteva riequipaggiarsi nel corso di una battuta di caccia avvistando una preda inattesa.

In effetti, la capacità di riequipaggiarsi durante la caccia può benissimo essere stato il principale vantaggio della tecnologia dei microliti. Questa tecnologia era impiegata probabilmente in situazioni nelle quali i cacciatori partivano alla ricerca di cibo senza avere in mente un tipo specifico di preda. Quando ne avvistavano una, potevano dare in breve alle piccole lame di pietra una nuova forma più adatta alla cattura di quella preda. Così la tecnologia dei microliti, che richiedeva un investimento di lavoro relativamente piccolo, era particolarmente adatta a situazioni in cui tipo e quantità di selvaggina erano imprevedibili. Di contro, la tecnologia specializzata, che richiedeva un considerevole investimento di tempo lavorativo, era più adeguata a situazioni in cui le risorse ricorrevano secondo uno schema prevedibile.

Benché gli utensili specializzati e quelli di uso più generalizzato funzionassero sulla base di principi diversi, non si deve pensare che fossero in contraddizione. Pare anzi che si integrassero bene. La strategia della diversificazione (messa in atto con la tecnologia dei microliti) riduceva i rischi della strategia specializzata. Combinando

La punta di lancia in corno a spine laterali era un utensile specializzato. Questo esemplare fu trovato a Star Carr, nell'Inghilterra settentrionale, che all'inizio del periodo postglaciale fu la sede di una comunità della cultura maglemosiana. I suoi membri potrebbero averlo usato per cacciare alci o cervi, fra altra selvaggina di grossa taglia. Nella Valle di Pickering, nello Yorkshire, dove si trova Star Carr, la stagione più favorevole per la caccia alle due specie era probabilmente l'inverno. Questa stagionalità faceva dei cervi e degli alci i candidati numero uno alla cattura con utensili specializzati. Questi utensili, che richiedono una preparazione considerevole, vengono usati con la massima efficienza nello sfruttamento di risorse che appaiono in epoca prevedibile, ma per periodi di tempo brevi, e che offrono potenzialmente alte rese. La punta fa parte della collezione del Rotunda Museum a Scarborough in Inghilterra.

i due metodi, un gruppo postglaciale si sarebbe potuto assicurare un rifornimento di cibo adeguato e ragionevolmente sicuro.

Una conferma di questa ipotesi è venuta da molti siti postglaciali nei quali sono stati trovati assieme utensili specializzati e microliti. Un esempio è costituito dal sito di Star Carr, nell'Inghilterra nordorientale, che fu abitato da una comunità maglemosiana. Negli anni cinquanta, Grahame Clark, dell'Università di Cambridge, vi scoprì circa 200 punte di lancia in corno con spine laterali accanto a utensili microlitici. Una interpretazione plausibile è che tali punte venissero usate per uccidere selvaggina di grossa taglia, come alci o cervi, nella stagione della loro maggior concentrazione, mentre le armi microlitiche dovevano essere usate per spedizioni di caccia nelle quali non si poteva prevedere in quali prede ci si sarebbe imbattuti; e ci si sarebbe potuti imbattere in una qualunque di una grande varietà di specie.

L'aver amalgamato tattiche specializzate e diversificate contribuì notevolmente a smorzare le fluttuazioni delle risorse, proprie dell'ambiente forestale. Forse la prova migliore che la disponibilità di cibo era migliorata proviene dai siti residenziali postglaciali, che spesso sembrano essere altrettanto permanenti e complessi quanto i primi villaggi agricoli. Pare in effetti che alcuni, come Sarnate in Lettonia, fossero occupati per tutto l'anno. In altri sono stati trovati i resti di abitazioni con pavimenti, focolari interni, attrezzature per la conservazione di cibi e partizioni per suddividere lo spazio interno. Strutture come queste sono molto lontane dal quadro abituale di un genere di vita fondato sulla caccia e raccolta. La loro relativa permanenza induce a pensare che nella vita di alcuni gruppi dediti alla caccia e raccolta sia avvenuto un mutamento fondamentale. Anziché spostarsi costantemente alla ricerca di cibo, ora il gruppo soggiornava in un luogo per gran parte dell'anno (se non per tutto l'anno), inviando piccole bande nel territorio circostante per eseguire compiti specifici di caccia e raccolta.

Alla prolungata permanenza nei luoghi residenziali si accompagnarono sviluppi sociali che innalzarono, infine, i gruppi di cacciatori-raccoglitori al di sopra del livello sociale rudimentale generalmente attribuito alle comunità di questo tipo. Emersero, per esempio, distinzioni di ricchezza, di potere e di rango fra i membri della comunità. La differenziazione sociale si osserva nel modo migliore nei cimiteri dell'epoca postglaciale, una ventina dei quali sono stati riportati in luce in Europa. Il più grande e più ricco è l'Oleneostrovskij Mogil'nik (Tumulo dell'Isola dei cervi), sul Lago Onega nella Russia settentrionale. Questo sito, che risale al VI millennio a.C., fu scavato negli anni cinquanta da Nina Nikolaevna Gurina dell'Istituto di archeologia di Leningrado. Recentemente John M. O'Shea dell'Università del Michigan e io abbiamo analizzato i reperti provenienti da esso per vedere che cosa potessero dirci sulla differenziazione sociale.

O'Shea e io eravamo particolarmente interessati alla distribuzione dei manufatti fra le tombe. Volevamo sapere se la distribuzione fosse abbastanza diseguale e sistematica da indicare l'esistenza di gruppi di ranghi all'interno della società. Abbiamo scoperto che il rango variava lungo varie dimensioni sociali. La prima dimensione era la ricchezza. Sulla base degli ornamenti recuperati nelle tombe è stato possibile distinguere tre gruppi. Molti degli ornamenti sono collane fatte di denti di animali e il tipo di denti può servire come indice della ricchezza del defunto. Le tombe dei membri del gruppo più ricco contenevano collane fatte con il bene più prezioso: denti d'orso. I membri del gruppo intermedio venivano sepolti con collane di incisivi di castoro o di denti d'alce. Nelle tombe dei membri del gruppo più povero non è stato trovato alcun ornamento.

Almeno altre due forme di differenziazione sociale sono emerse nelle tombe dell'Isola dei cervi. Una è l'esistenza di ruoli speciali all'interno della comunità. Un gruppo di tombe, occupate esclusivamente da scheletri maschili, conteneva ricchi depositi di punte di lancia in osso, ma nessun altro manufatto; sembrerebbe trattarsi di resti di uomini che avevano speciali responsabilità come cacciatori, conferite loro dalla comunità. Un altro gruppo di tombe, con un'abbondanza insolita di ornamenti, potrebbe essere appartenuto a specialisti del rituale. A differenza dei membri ordinari della comunità (sepolti sdraiati in tombe orizzontali), questi presunti specialisti del rituale venivano inumati in tombe a pozzo verticali. Pare inoltre che qualche distinzione di ricchezza e di rango venisse trasmessa da una generazione all'altra, come dimostra la presenza di tombe con imponenti suppellettili funerarie assieme a scheletri di bambini.

La conclusione che le società di cacciatori-raccoglitori postglaciali siano passate per un processo di differenziazione sociale che le innalzò al di sopra del livello egualitario è rafforzata da prove che emergono in modo generalizzato dall'intera zona a foreste. Le foreste settentrionali sono piene di oggetti intagliati in legno, osso o pietra, risalenti al periodo postglaciale. I soggetti raffigurati più spesso sono animali forestali, come orsi o alci, o uccelli acquatici; si trovano anche figurine antropomorfiche. Un aspetto interessante di questi oggetti è che, con il progredire del periodo postglaciale, ne furono prodotti in numero sempre maggiore ed essi risultano distribuiti in tutte le regioni forestali.

Per la maggior parte questi oggetti a intaglio non avevano alcuna chiara applicazione utilitaristica. Pare che avessero, invece, una funzione sociale: in una cultura priva di denaro, oggetti ben intagliati, la cui produzione richiedeva molto tempo e una considerevole abilità, assolvevano la funzione di contrassegni di valore. Il loro possesso suscitava considerazione in altri membri del gruppo. Barbara Bender, dello

I microliti sono frammenti di pietra che, dopo essere stati lavorati in modo da munirli di una punta aguzza ed essere stati affilati, venivano montati su un manico d'osso, di legno o di corno. Quelli qui raffigurati provengono dal sito di Deercar, nello Yorkshire, in Inghilterra, e furono prodotti nell'VIII millennio a.C. o all'inizio del VII millennio. Queste microlame, che si trovano in quasi tutti i siti postglaciali, furono la componente principale di una cronologia polivalente. Avvistando una preda, un cacciatore poteva togliere il microlito dal manico e scheggiarlo in modo da modificarne la forma; poteva così riequipaggiarsi per l'uccisione di una grande varietà di prede. La tecnologia dei microliti veniva usata probabilmente quando una pattuglia di cacciatori partiva senza avere in mente una preda particolare. Essa integrava la tecnologia specializzata, che era adatta a spedizioni progettate accuratamente. Questi microliti appartengono alla collezione del City Museum di Sheffield in Inghilterra.

University College di Londra, ha sottolineato (in un contesto sociale lievemente diverso) che la diffusione di contrassegni di valore potrebbe riflettere una maggiore competizione sociale. O'Shea e il mio collega Paul Halstead hanno suggerito che forse i contrassegni assolvevano anche la funzione di «accumulazione sociale», venendo ammassati in periodi di abbondanza per essere poi scambiati con cibo o altri beni di prima necessità in tempi di bisogno. Così, per un individuo essi potevano non solo esprimere ricchezza e rango, ma rappresentare anche un'assicurazione aggiuntiva contro fluttuazioni nella disponibilità di cibo.

L'emergere di una gerarchia sociale completa il quadro dei gruppi di cacciatori-raccoglitori postglaciali, che ho fin qui delineato. La differenziazione sociale (e il concomitante aumento della competizione sociale) si andò ad aggiungere alla combinazione di tecnologia specializzata e di tecnologia polivalente, all'uso di conservare i cibi e all'installazione di insediamenti permanenti o quasi. Una tale costellazione costituisce un adattamento complesso all'ecologia della zona a foreste eurasiatica. Secondo me l'adattamento complesso raggiunto dai cacciatori-raccoglitori dovrebbe servire come caratteristica essenziale alla definizione del Mesolitico. Il Mesolitico è l'intervallo di tempo compreso tra la fine delle culture del Paleolitico superiore, durante la glaciazione pleistocenica, e l'introduzione dell'agricoltura, che segna l'inizio del Neolitico. Esso viene usato talvolta o come una definizione cronologica o come una definizione economica relativamente semplice, che lo caratterizza come periodo postglaciale di caccia e raccolta. Nessuna delle due definizioni coglie, tuttavia, appieno i risultati conseguiti dallo sviluppo sociale del Mesolitico e rappresentati dal complesso adattamento delle tecniche di caccia e raccolta.

In considerazione del grande successo conseguito da questo adattamento nello stabilizzare la disponibilità di cibo, sorge naturale chiedersi perché esso abbia infine ceduto il passo alla diffusione dell'agricoltura in tutte le principali regioni d'Europa. Secondo l'opinione tradizionale, l'agricoltura si sarebbe propagata rapidamente e uniformemente a partire dal Vicino Oriente, introdottavi da emigranti o diffusasi per imitazione culturale. Recentemente Albert J. Ammerman, della Princeton University, e Luigi L. Cavalli-Sforza, della Stanford University, hanno costruito un modello matematico fondato su tali assunti. Le conoscenze sulla diffusione dell'agricoltura nell'Europa sudorientale e centrale si conciliano molto bene con esso. Nell'Europa settentrionale e orientale, invece, il modello non risulta altrettanto calzante.

Lungo la costa atlantica settentrionale e nelle foreste dell'Europa orientale e settentrionale, l'agricoltura si propagò molto più lentamente di quanto avrebbe previsto il modello. Appare infatti che, nel complesso, il progredire dell'agricoltura non assomigli per nulla a una marcia costante. Sembra invece che sia stato più simile a una serie di brevi balzi in avanti, interrotti da lunghi periodi di attesa. Per esempio, l'agricoltura era praticata nella Germania settentrionale attorno al 4500 a.C. Vi fu poi una lunga stasi e solo 1200 anni dopo le attività agricole comparvero anche in Danimarca e nella Svezia meridionale. Secondo il modello, l'agricoltura avrebbe dovuto prendere piede nella zona a foreste dell'Europa orientale fra il 5000 e il 4000 a.C., mentre in realtà fu adottata due-tre millenni dopo. Essa si diffuse particolarmente tardi in regioni come l'area del Baltico orientale, la Spagna settentrionale e l'Ucraina settentrionale, che sono ricche di risorse acquatiche. È degno di nota che questi ambienti sono appunto quelli nei quali fu più sviluppato il complesso adattamento delle tecniche di caccia e raccolta.

Quando si esamina la documentazione archeologica su una scala geografica minore, diventa chiaro che, in alcune parti della zona a foreste, l'agricoltura fu adottata ancora più tardi di quanto suggerirebbe un'indagine generalizzata sull'Europa. Alcuni studiosi sono stati tentati di considerare il primo osso di un animale domesticato o il primo chicco di un cereale trovato in un sito preistorico come una prova della pratica, in quel sito, dell'agricoltura. Tali generi potrebbero però essere stati ottenuti per mezzo di scambi commerciali. Ha molto più senso considerare la comparsa costante di specie domesticate come criterio per il passaggio dalla caccia e raccolta all'agricoltura. In alcune aree la differenza fra le due date è considerevole. In Finlandia, nella regione del Baltico orientale e in molte parti dell'Unione Sovietica resti isolati di specie domesticate furono depositati anche mille anni prima dell'inizio di un'agricoltura sistematica.

Quel che emerge da un esame della zona di frontiera che divide gli agricoltori dai cacciatori-raccoglitori è quindi l'immagine di un lungo periodo in cui l'agricoltura era già accessibile ai cacciatori-raccoglitori, i quali però non avevano ancora cominciato a metterla in pratica. Questo fa sorgere la domanda del perché i cacciatori-raccoglitori siano stati così lenti nell'adottare le tecniche agricole. Io vedo la ragione di questo ritardo nel fatto che l'agricoltura non era necessariamente vantaggiosa, particolarmente per le comunità specializzate nello sfruttamento delle risorse acquatiche.

In queste culture l'adozione dell'agricoltura avrebbe fatto perdere un grande investimento in tecnologia specializzata e in forme di organizzazione sociale adattate specificamente a un sistema complesso di caccia e raccolta. L'introduzione dell'agricoltura avrebbe causato inoltre problemi di programmazione di tempi, poiché nella zona a foreste eurasiatica, la stagione per la raccolta delle specie vegetali è perlopiù l'autunno, che è anche il periodo più propizio alla caccia. In ogni caso l'incentivo per l'adozione dell'agricoltura dev'essere stato scarso, almeno inizialmente, non essendovi un bisogno pressante di aumentare la produttività. La strategia specializzata per la caccia e raccolta permetteva di sostenere densità di popolazione relativamente elevate e di assorbire una certa fluttuazione delle risorse. La spiegazione del passaggio all'agricoltura non sta, quindi, nella sua superiorità intrinseca.

Una spiegazione plausibile è che l'agricoltura fu introdotta per ovviare alla mancanza di risorse specifiche. L'essenza della complessa economia di caccia e raccolta era l'equilibrio fra le poche specie che venivano sfruttate intensivamente per mezzo di una tecnologia specializzata e il gruppo maggiore di prede «incontrate casualmente» e catturate con utensili polivalenti. Finché questo delicato equilibrio fu mantenuto, coloro che mettevano in pratica un sistema complesso di caccia e raccolta riuscirono a ridurre i rischi e a mantenere la produttività. La scomparsa di una singola risorsa sfruttata in modo intensivo poteva però distruggere l'equilibrio e mettere in crisi la comunità di cacciatori-raccoglitori. Varie aree d'Europa hanno fornito prove di crisi di questo tipo. In Danimarca scomparvero le ostriche; nella Finlandia meridionale le foche; nella Russia nordoccidentale le castagne d'acqua assieme ai bovini selvatici e ai cinghiali. Pare che in tutt'e tre queste regioni l'agricoltura abbia attecchito come ripiego per rimediare alla crisi causata dal venir meno di determinate risorse.

L'integrità di economie fondate sulla caccia e raccolta potrebbe anche essere stata infranta dalla competizione con i gruppi che praticavano l'agricoltura. Nella zona di frontiera comunità che praticavano la caccia e raccolta e comunità agricole devono essersi trovate in concorrenza per una varietà di risorse, fra le più significative delle quali era il territorio. La fine della cultura della ceramica ad alveoli è un esempio di come la perdita di territorio potesse portare con sé gravi conseguenze per popolazioni di cacciatori-raccoglitori.

Attorno al 2600 a.C. gruppi appartenenti a quella cultura occupavano l'interno della Svezia centrale oltre alle aree costiere e alle isole fra la Svezia e la Finlandia. Tale vasto territorio forniva una varietà di selvaggina terrestre, oltre alle foche, nella cui caccia quei gruppi erano specializzati. Non più tardi del 2300 a.C. nell'interno della Svezia fu introdotta l'agricoltura. Di conseguenza, le comunità della ceramica ad alveoli videro il loro territorio limitarsi alla costa e aumentare la loro dipendenza dalle foche. Una volta distrutta la loro economia equilibrata, esse si trovarono male equipaggiate per far fronte alle fluttuazioni nella popolazione di foche che, tra il 2000 e il 1800 a.C., si accompagnarono a un mutamento di clima. L'economia della caccia e raccolta scomparve ben presto, e la popolazione della ceramica ad alveoli fu probabilmente assorbita da quella agricola dell'interno.

Un'altra forma di competizione potrebbe aver avuto un ruolo nella transizione all'agricoltura: quella all'interno di una società di cacciatori-raccoglitori. Come abbiamo visto, nel tardo Mesolitico sarebbero aumentate la differenziazione sociale e la lotta per conseguire un rango, e buona parte di questa lotta fu orientata verso oggetti che avevano la funzione di contrassegni di valore. Nella zona di frontiera in cui comunità di agricoltori e di cacciatori-raccoglitori si trovarono in contatto, questi ultimi dovettero venire a conoscenza di oggetti che non conoscevano e ai quali attribuirono forse un grande valore. Per esempio, nel III millennio a.C. i commerci nell'ambito di una catena di comunità agricole potrebbero aver portato ceramica ben cotta o asce ben levigate dall'Europa centrale verso nord fino alla Scandinavia. I cacciatori-raccoglitori di quest'area potrebbero aver visto quegli oggetti e averli desiderati come simboli di un certo rango. Per ottenerli, potrebbero avere intensificato la ricerca delle merci da offrire in cambio, per esempio pelli di castoro. Ben presto, però, dovettero esaurire le disponibilità o raggiungere il punto in cui un'ulteriore intensificazione della caccia sarebbe diventata antieconomica. In casi del genere potrebbero essersi convertiti alla maggiore capacità produttiva dell'agricoltura come mezzo per soddisfare il desiderio di acquisizione.

Come spero sia stato chiarito dai paragrafi precedenti, non penso che la transizione all'agricoltura abbia avuto luogo dappertutto, nella zona a foreste, per la medesima ragione. Al contrario, in ogni regione potrebbe avere operato uno o l'altro di vari meccanismi e, da una regione all'altra, le cause immediate possono essere state diverse. Ne segue che il tentativo di trovare una spiegazione singola, universale, per l'adozione dell'agricoltura è destinato all'insuccesso. A mio giudizio un programma di ricerca più fecondo sarebbe quello che consiste nell'esaminare le circostanze specifiche della transizione in molte aree minori.

Un tale esame potrebbe modificare sotto vari aspetti la nostra immagine dei cacciatori-raccoglitori. Penso che si troverebbe che si è sottovalutata la persistenza dell'adattamento della caccia e raccolta. Si potrebbe inoltre correggere la nozione che la sostituzione della caccia e raccolta con l'agricoltura abbia costituito un corso di progresso lineare. Nella cronologia e nella funzione, lo sviluppo degli adattamenti complessi di caccia e raccolta fu parallelo a quello dell'agricoltura. Prima o poi questo parallelismo sarà riconosciuto e le due forme saranno viste per quello che sono in realtà: tentativi di ridurre i rischi e di aumentare la produttività dell'ambiente intensificando il lavoro nel territorio lasciato libero dal ritiro dei ghiacciai.

La diffusione dell'agricoltura dal Vicino Oriente all'Europa subì un rallentamento considerevole nella zona a foreste. Su ciascuna di queste cartine l'area in colore più intenso corrisponde al territorio in cui aveva preso piede l'agricoltura, mentre l'area in grigio corrisponde al territorio occupato dai cacciatori-raccoglitori. In colore chiaro è la «zona di disponibilità», dove gruppi di cacciatori-raccoglitori avevano avuto contatti con agricoltori e avevano appreso le tecniche dell'agricoltura, ma non le avevano ancora adottate. Nella regione dell'Atlantico settentrionale, comprese Danimarca e regioni limitrofe, l'agricoltura era disponibile già attorno al 4500 a.C., ma non fu adottata sino a poco prima del 3000 a.C. La sua conoscenza si diffuse nell'Europa orientale a nord del Mar Nero fra il 5000 e il 4000 a.C., ma le tecniche non furono introdotte sino al 2500 a.C. o più tardi. In queste regioni fiorirono le complesse culture dedite alla caccia e alla raccolta. Il lungo ritardo con cui l'agricoltura fu adottata induce a pensare che non fosse necessariamente vantaggiosa per gruppi che avevano subito un adattamento così complesso.

Darwin geologo

Oggi si tende più a considerarlo un biologo, ma egli si riteneva un geologo e, nei cinque anni passati sulla Beagle, *la sua principale attività di ricerca riguardò la geologia, a cui i suoi lavori diedero un importante contributo*

di Sandra Herbert

Oggi si tende ad annoverare Darwin essenzialmente tra i biologi, interessato - come egli era - all'origine e all'evoluzione delle specie animali e vegetali. Ma è piuttosto come geologo che egli prese parte al famoso viaggio della *Beagle*, durato dal 1831 al 1836, e negli anni immediatamente successivi a quelli del viaggio egli si considerò sempre un geologo. I suoi contributi alla geologia furono fondamentali. Le sue meticolose osservazioni sul campo, effettuate durante il viaggio, fornirono raccolte di materiale che rimangono preziose. Le sue intuizioni sull'origine delle scogliere coralline posero le basi per la concezione odierna della materia. Infine, fu in gran parte la geologia a condurlo alle sue concezioni sull'evoluzione, o «trasmutazione», come voleva il vocabolario dell'epoca.

L'interesse di Darwin per la geologia, manifestatosi durante il viaggio, non derivava da una frequentazione di lunga data della materia. Anzi, all'inizio egli era pronto più che altro a raccogliere insetti e altri invertebrati: sin da ragazzo, infatti, era stato un entomologo dilettante, e nei due anni trascorsi come studente di medicina all'Università di Edimburgo aveva ricevuto una buona preparazione allo studio degli invertebrati.

Durante il viaggio, la geologia passò al primo posto per varie ragioni. Innanzitutto l'Ammiragliato aveva bisogno, per quel viaggio in particolare, di un esperto in geologia, e la cosa deve per forza aver influito sul modo in cui Darwin, nei colloqui preliminari, si era presentato a Robert FitzRoy, comandante della nave. FitzRoy era consapevole degli interessi economici serviti dalla geologia. Nel 1830, dopo una missione nell'America Meridionale con la *Beagle*, egli aveva fatto notare, per esempio, che un geologo o un mineralogista sarebbe stato in grado di accertare se, nelle zone montuose della Terra del Fuoco, vi erano davvero giacimenti metalliferi. Analogamente, quando venne il momento di allestire la *Beagle* per il suo secondo viaggio, Francis Beaufort, l'idrografo della Marina britannica, mise in rilievo l'importanza di una ricerca geologica sulle isole coralline. Egli richiamò l'attenzione sul fatto che «di quando in quando, esse offrono ottimi porti naturali chiusi da terre» e fece notare che al momento esse erano di grande interesse per i geologi.

La seconda ragione che portò Darwin verso la geologia durante il viaggio della *Beagle* aveva a che fare con lo stato particolarmente allettante del settore a quel tempo. Gli anni trenta del secolo scorso furono un decennio d'oro per la geologia: i geologi definirono i sistemi cambriano, siluriano e devoniano, completando così su scala globale la successione degli strati geologici. Furono anche il decennio in cui la maggior parte dei geologi inglesi abbandonò il tentativo di sincronizzare la storia biblica con quella geologica.

La persona che mise Darwin in contatto con i nuovi sviluppi in questo campo fu John Stevens Henslow, professore di botanica all'Università di Cambridge (a cui Darwin si era iscritto dopo aver lasciato Edimburgo), membro della Geological Society di Londra e mentore di Darwin. Nel 1831, Darwin, sollecitato da lui, iniziò per la prima volta uno studio serio della materia. Fino a quel momento aveva evitato l'argomento, essendo rimasto deluso dalle lezioni di geologia tenute da Robert Jameson a Edimburgo.

Nell'estate del 1831 Henslow fece in modo che Darwin svolgesse osservazioni sul campo con Adam Sedgwick, professore di geologia a Cambridge. Questo lavoro indusse Sedgwick a definire il sistema cambriano. Fu così che, solo qualche mese prima di imbarcarsi sulla *Beagle*, Darwin fu addestrato dal più illustre esponente inglese dell'osservazione diretta della natura, a quel tempo impegnato nella sua opera di maggior rilievo, nei metodi di ricerca sul campo. Di ritorno dalle escursioni con Sedgwick, Darwin trovò ad attenderlo la lettera con l'offerta di partecipare alla spedizione della *Beagle*.

La terza ragione per la quale Darwin scelse la geologia dipendeva certamente dal fatto che essa si adattava alla sua personalità. Negli anni trenta del secolo scorso la geologia univa, al tradizionale rilievo dato dalla storia naturale alla peculiarità delle cose, un'apertura alla teorizzazione: una combinazione che si adattava assai bene alle capacità e ai gusti di Darwin. In una breve nota autobiografica scritta poco dopo il viaggio, Darwin rammenta un desiderio dell'infanzia: «poter saper qualcosa su ogni ciottolo che si trova davanti al portone di casa: si trattava della mia prima e unica aspirazione geologica a quel tempo». Il suo talento per l'astrazione lo indusse a ricordare il proprio interesse per «la registrazione della stratificazione e della natura delle rocce e dei fossili in molti punti, sempre ragionando e prevedendo quello che si sarebbe trovato altrove».

La decisione di Darwin di concentrarsi sulla geologia maturò agli inizi del viaggio. Ciò che sembra aver portato all'unione delle forze che premevano dall'interno e dall'esterno fu la prospettiva di scrivere qualche cosa di proprio. Mentre si preparava per il viaggio egli notò di aver paura di «essere sopraffatto dal numero degli argomenti di cui dovrò occuparmi». Mentre la *Beagle* era all'ancora a Santiago, una delle isole dell'arcipelago del Capo Verde, gli venne in mente un piano che ebbe l'effetto di organizzare le sue attività per il resto del viaggio: avrebbe scritto un libro sulla geologia dei luoghi visitati. Nella sua autobiografia ricorda che, dopo aver dedotto la storia geologica dell'isola di Santiago studiandone le rocce, «mi venne in mente per la prima volta che avrei anche potuto scrivere un libro sulla geologia dei vari paesi visitati, e questo mi fece fremere di piacere».

Che in così giovane età - avrebbe compiuto fra breve 23 anni - pensasse a se stesso come autore di un'opera geologica di rilievo è certo segno di una grande ambizione, la quale peraltro era molto realistica, poiché Darwin sapeva che la geologia dell'America Meridionale, specie delle estremità meridionali del continente, era stata così poco esplorata che il suo lavoro non avrebbe potuto non essere originale. Egli inoltre subiva l'influenza e l'ispirazione di due capolavori letti poco tempo prima: *Personal Narrative* di Alexander von Humboldt, sui suoi viaggi nell'America Meridionale, opera di impronta fortemente geologica, ma che riguardava soprattutto la parte settentrionale del continente, e *Principles of Geology* di Charles Lyell, il cui primo volume era uscito nel 1830 e piacque a Darwin per lo stile chiaro e il contenuto stimolante.

Con in mente la futura pubblicazione, Darwin partì dalle isole del Capo Verde per raccogliere un numero esauriente di campioni geologici e fare osservazioni il più possibile complete. «Ho cercato di procurarmi esemplari di ogni varietà di roccia e ho scritto note su tutto.» Il lavoro portò a una pubblicazione in tre parti sulle scoperte geologiche del viaggio: *The Structure and Distribution of Coral Reefs* (1842), *Geological Observations on the Volcanic Island Visited during the Voyage of H. M. S. Beagle, together with some Notices of the Geology of Australia and the Cape of Good Hope* (1844) e *Geological Observations on South America* (1846).

Raccogliendo campioni durante il viaggio e descrivendo per esteso le proprie scoperte, Darwin diventò geologo. È in questa luce che egli si vedeva ed è sempre in questa luce che lo vedevano gli altri. Alla fine della spedizione, parlando di sé, così scriveva su un taccuino personale: «Io, geologo». La sua identità si era modellata attorno alla sua ambizione.

Era tipico in lui che l'ambizione agisse a più di un livello. Ogni giorno egli perseguiva il proprio intento originale di raccogliere e di osservare sistematicamente. E gli tornavano molto utili le sue abitudini metodiche. Da sportivo aveva preso nota della selvaggina; da geologo fu altrettanto diligente nel numerare i propri esemplari, registrarne la localizzazione e continuare diligentemente a prendere appunti. A tutto un altro livello rifletté sul significato delle proprie scoperte e sviluppò le ipotesi che hanno portato alla sua fama duratura di pensatore profondamente originale.

Le note riportate dal viaggio, che si trovano nella biblioteca di Cambridge, rivelano le sue abitudini quotidiane di lavoro. A differenza di altri ricercatori inglesi, che spesso ritornavano ogni estate nello stesso sito, Darwin fu costretto a raccogliere tutto ciò che voleva in una sola volta. Gli appunti presi durante il viaggio sulla sua prima escursione geologica illustrano questo metodo. Il 16 gennaio 1832 la nave gettò l'ancora nel porto di Praia, nell'isola di Santiago, e le sue note del 17 e 18 gennaio riguardano la Quail Island (ora Ilhéu de Santa Maria): «La Quail Island, un puntino desolato situato vicino a Porto Praia. La sua forma è ovale da N. a S. e la sua circonferenza è a mala pena di un miglio. Intorno a Porto Praia vi sono parecchie colline a tronco di cono; l'isola può essere considerata una di quelle colline, con ai piedi il mare invece di una valle

L'isola di Moorea, nel Pacifico meridionale, vicino a Tahiti, influenzò il pensiero di Darwin sull'origine delle scogliere coralline. Nel 1835, in una sosta della *Beagle* a Tahiti, da una collina Darwin vide pressappoco questo panorama di Moorea (allora Eimeo), a circa 25 chilometri di distanza, e ipotizzò che se l'isola, circondata da scogliere coralline, fosse sprofondata lentamente, al suo posto si sarebbe formato un atollo corallino.

Nella veduta di Moorea, a livello del mare, vicino al lato prospiciente Tahiti, è visibile parte della scogliera corallina. È l'isoletta coperta di alberi al centro. Darwin estese la propria teoria sulla formazione delle scogliere e vi incluse le scogliere costiere, le barriere coralline e gli atolli.

sabbiosa. - L'erosione da parte del mare attorno alla sua base ne mette alla scoperto una buona sezione e, studiando attentamente l'isola, ho pensato che forse avrei potuto trovare una valida chiave di volta per il paese circostante. - Incomincerò con gli strati inferiori e descriverò minuziosamente tutta la sequenza.»

La descrizione particolareggiata degli strati era accompagnata da campioni di roccia. Per esaminare i campioni Darwin si serviva di attrezzi semplici: una lente d'ingrandimento, una bottiglia di acido per analizzare l'alcalinità, un cannello ossidrico per verificare la reazione al calore di ogni campione, un goniometro per misurare gli angoli e una calamita. Egli stava molto attento a etichettare ogni campione con un numero che fosse poi lo stesso riportato sia nell'elenco dei campioni sia nelle sue note. Quindi tutta la sua collezione era ed è ancora utilizzabile, a differenza di molte altre raccolte non fruibili per la povertà dei dati che le accompagnano.

Oltre a descrivere gli strati e a raccogliere campioni di rocce, Darwin avanzava congetture sull'origine di quello che vedeva. Sulla Quail Island, per esempio, concludeva: «Nel complesso riterrei probabile [aveva cancellato *evidente*] che sugli strati marini, dopo che si sono depositati sulle rocce vulcaniche sottostanti, si sia sparso uno strato di materia fusa: che tutta la massa si sia sollevata; dopo di che, oppure nel momento stesso, vi è stato un *parziale sprofondamento*.»

Nonostante le capacità di raccoglitore e di osservatore, la fama di Darwin come geologo si basa sulle sue teorizzazioni. Ciò che lo distingueva dalla maggior parte dei suoi colleghi era la capacità di trattare sia questioni astratte di struttura e di causalità, sia questioni di minuziosa osservazione, passando con facilità dall'uno all'altro livello. Il suo impulso a teorizzare, accompagnato da una buona disposizione ad affrontare problemi controversi, si rivelò fin dalle primissime note sulla Quail Island. Inoltre, la sua ambizione di geologo crebbe via via che, dai due anni previsti inizialmente, il viaggio si prolungò fino a raggiungere i cinque anni. Avendo come guida i testi di geologia dell'epoca, Darwin si concesse di diventare più speculativo nel modo di affrontare l'argomento.

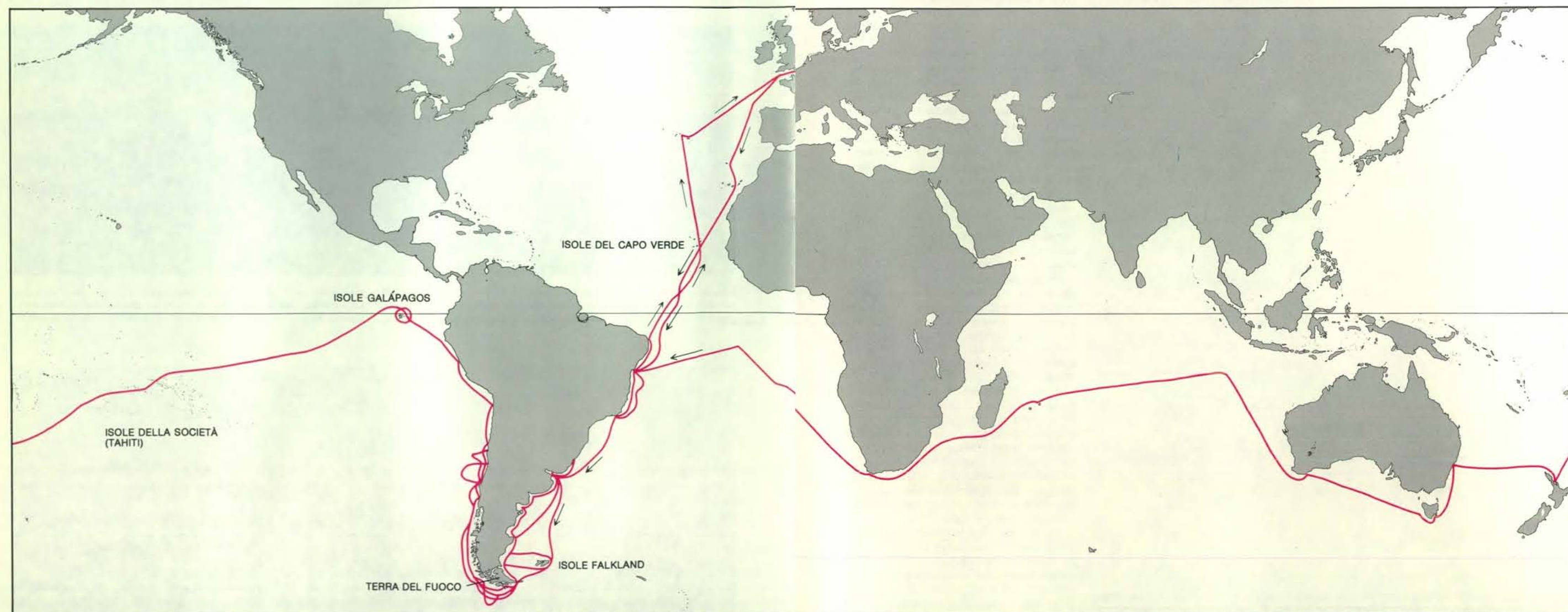
L'evoluzione di Darwin come teorico si può seguire nel modo migliore esaminando i due temi di più vasta portata da lui perseguiti: i movimenti verticali della crosta terrestre e la trasmutazione, ovvero l'evoluzione delle specie. Il primo tema derivava dalle letture fatte: egli fece sua un'idea dei suoi predecessori e la sviluppò oltre le dimensioni che questi le avevano dato. Anche l'altro tema, perseguito in un secondo tempo durante la spedizione, e in maniera più tortuosa, gli fu suggerito dalla lettura dei predecessori, ma in questo caso dovette vincere la loro influenza, perché la maggior parte si era espressa negativamente nei confronti dell'evoluzione.

Alla fine del viaggio il primo tema era ben chiaro nella mente di Darwin, che nel 1836, nell'ultima tappa della spedizione, annotò sul suo «Taccuino rosso»: «La geologia di tutto il mondo risulterà semplice.» Egli alludeva alla nozione che la geologia mondiale era governata dal semplice movimento verso l'alto e verso il basso, cioè dall'innalzamento e dallo sprofondamento (o subsidenza), della superficie terrestre. A suo parere l'equilibrio di queste due forze era la chiave per stabilire le caratteristiche principali della superficie della Terra.

Riprendendo questo tema, Darwin continuava un lavoro iniziato da altri geologi. L'unica persona il cui pensiero lo influenzò molto profondamente fu Lyell. Seguendo una tradizione instaurata verso la fine del primo decennio del secolo scorso dal geologo scozzese James Hutton, Lyell sosteneva in *Principles of Geology* che la superficie terrestre è in perpetuo movimento. Come la maggior parte dei geologi suoi contemporanei, egli credeva che presupporre mutamenti del livello della superficie fosse un modo più semplice di spiegare la presenza di resti marini nei continenti che non presupporre mutamenti su scala mondiale del livello del mare. Ma, diversamente da molti suoi pari, riteneva che i mutamenti del livello della superficie avessero luogo gradualmente. Egli era contrario all'opinione di molti geologi, secondo i quali la Terra attraversava periodi di stabilità seguiti da periodi di innalzamento.

Darwin affrontò i fenomeni osservati durante il viaggio tenendo presenti le opinioni di Lyell. Mentre le osservazioni sul campo, fatte da Lyell, si riferivano a zone dell'Europa abbastanza delimitate, Darwin svolse le sue osservazioni in vari oceani e continenti. Le sue differenti esperienze incisero sul modo in cui applicò le idee di Lyell. Il cambiamento di maggior rilievo che apportò al concetto di innalzamento e di subsidenza espresso da Lyell fu quello di applicarlo su scala più vasta, la scala dei continenti, e non a un'area delimitata.

Quando esplorò quella che oggi è l'Argentina, Darwin fu colpito dall'immensità delle pianure e dalla persistenza per grandi tratti lungo la costa di litorali rialzati di altezza quasi identica. Da queste osservazioni dedusse che era in atto una forza di sollevamento che agiva su una vasta area geografica e tanto lentamente da non turbare la continuità dell'altezza delle pianure e dei litorali rialzati. Raggiunte le Ande, rimase colpito dalle loro grandi dimensioni rispetto alle montagne dell'Europa. Vicino a Valparaiso, in Cile, trovò le prove di una linea di costa in sollevamento: la presenza di conchiglie marine molto all'interno e al di sopra del livello costiero di acqua alta. Il 20 febbraio 1835 assistette a Valparaiso a un terremoto accompagnato da un sollevamento della superficie terrestre di varie decine di centimetri.

Tutte queste osservazioni insieme contribuirono a portarlo alla conclusione che l'intero continente sudamericano si stava sollevando e che la formazione delle Ande

Il percorso della *Beagle* dal 1831 al 1836, con Darwin a bordo in qualità di ricercatore scientifico, è ricostruito sulla base delle note tenute da Robert FitzRoy, il comandante della nave. In un libro su due dei suoi viaggi, uscito nel 1839, FitzRoy includeva una mappa analoga, da lui stesso definita «carta generale che presenta le principali rotte dell'H.M.S. *Beagle*». Poiché a FitzRoy interessavano soprattutto le «rotte principali», la mappa non presenta tutti gli spostamenti della nave in occasione di escursioni come quelle, numerose, da e per le isole Falkland.

e il verificarsi di terremoti e di eruzioni vulcaniche non erano, in quel continente, se non effetti collaterali di tale movimento.

Nella successiva applicazione della nozione di sollevamento e di subsidenza, espressa da Lyell, Darwin diede alla geologia il suo contributo più duraturo, al di fuori di quello dato alla teoria dell'evoluzione. Si tratta della teoria sull'origine delle scogliere coralline, pubblicata per intero nel 1842 con il titolo di *The Structure and Distribution of Coral Reefs*.

Potrà forse sembrare sconcertante il fatto che - come disse una volta il geologo H. W. Menard - «di quando in quando Darwin sembra interessato all'origine delle scogliere coralline solo nella misura in cui essa gli fornisce le prove della subsidenza regionale del fondo marino». Peraltro, tutte le perplessità scompaiono se ci si rende conto che Darwin formulò la sua teoria come un corollario dell'interpretazione del movimento della crosta terrestre. Per lui, come per Hutton e per Lyell, il sollevamento di un'area della superficie terrestre doveva essere controbilanciata dall'abbassamento di un'altra. Una volta giunto alla conclusione che tutta l'America Meridionale, e presumibilmente anche tutta l'America Settentrionale, si stava sollevando, egli cercò in qualche altro posto una corrispondente zona di subsidenza. «Questa area corrispondente - si chiedeva - non è forse compresa nella grande distesa del Pacifico settentrionale e meridionale?» Ancora una volta Lyell aveva dato a Darwin i punti fondamentali della sua argomentazione. In *Principles of Geology* egli avanzò l'ipotesi che l'oceano Pacifico fosse un'area di subsidenza e offrì anche una spiegazione delle forme di vari tipi di scogliere coralline. La sua tesi era che, siccome i coralli non crescono al di sopra della superficie dell'acqua e i coralli costruttori di scogliere crescono soltanto in acque poco profonde, le scogliere non possono formarsi che su piattaforme sottomarine preesistenti, come sommità di montagne e crateri vulcanici che si sono poi abbassati. Le isole coralline devono poggiare sulle prime, gli atolli sui secondi.

Darwin pensava che la spiegazione dell'origine delle scogliere coralline, data da Lyell, fosse ingegnosa, ma era convinto che Lyell non avesse tenuto conto a sufficienza della vastità dell'oceano Pacifico. A suo modo di vedere un'area così grande non poteva non avere subito una subsidenza altrettanto estesa. Era sbagliato quindi presupporre l'esistenza di montagne e vulcani poco sommersi per spiegare la presenza di scogliere coralline in mezzo all'oceano. A suo parere era meglio immaginare che le scogliere si formassero su piattaforme che esse stesse avevano costruito via via che il fondo dell'oceano si era abbassato sotto di loro. Secondo Darwin - e questa sua opinione ebbe una conferma definitiva soltanto in questo secolo, negli anni cinquanta - le scogliere coralline potevano avere un notevole spessore.

Che dire delle varie forme di scogliere? Qui fu istruttiva la prima osservazione di una scogliera corallina, effettuata da Darwin. Nel mese di novembre del 1835, la *Beagle* fece sosta a Tahiti. Darwin risalì una collina fino a un'altezza di 700-1000 metri e guardò verso la bellissima isola di Eimeo (l'odierna Moorea), completamente circondata da una scogliera. Mentre la vedeva spuntare dal suo «lago vetroso» (come scrisse più tardi), pensò: «Togli il gruppo centrale di montagne, e rimane un'isola con laguna.» Ma come poteva una montagna che si abbassava lasciare una scogliera a forma d'anello anziché un'isola a forma di disco? Perché, affermava Darwin, la laguna formatasi là dove un tempo si ergeva la montagna conteneva acqua stagnante ed era troppo poco profonda per offrire condizioni ambientali adatte alla crescita di coralli costruttori di scogliere.

Sulla via del ritorno in Inghilterra, Darwin generalizzò là sua intuizione. Divise le scogliere in tre categorie: le scogliere costiere o frangenti, le scogliere a barriera o barriere coralline e gli atolli. Le scogliere costiere (separate di poco dalla terraferma da acque poco profonde) si troverebbero là dove si sarebbero accresciute dal litorale sul lato verso il mare aperto. Secondo Darwin questo tipo di scogliera si formerebbe quando la superficie è stabile o in fase di sollevamento. Le barriere coralline (separate dalla terraferma da un canale profondo) si formerebbero quando si abbassa una scogliera costiera. I coralli crescono meglio in mezzo ai frangenti e il bordo esterno della scogliera costiera darebbe origine appunto alla barriera. Il canale si formerebbe a mano a mano che la superficie verso l'interno si abbassa. Un atollo si formerebbe quando la superficie sprofonda interamente, ma lentamente; altrimenti il corallo verrebbe sommerso.

Cosa interessante, la teoria di Darwin comportava la conclusione che una forma di scogliera si evolveva in un'altra, che la velocità del movimento del fondo oceanico era abbastanza graduale da non sommergere un buon numero di scogliere e che, notando la presenza di questa o di quella categoria di scogliere, era possibile forse individuare le aree del fondo oceanico che si innalzavano o si abbassavano. (Le scogliere costiere indicherebbero stabilità o sollevamento, le barriere e gli atolli abbassamento.) Anche se in seguito fu contestata in alcuni punti specifici, la teoria di Darwin sulla struttura e sulla distribuzione delle scogliere coralline continua a godere di grande considerazione. Giudicata come teoria, essa rimane un modello di semplicità e di potenza esplicativa.

DARWIN LYELL SEDGWICK HENSLOW

Gli scienziati raffigurati insieme a Darwin (come erano intorno al 1840) ebbero su quest'ultimo grande influenza in campo geologico. I *Principles of Geology* di Lyell furono per Darwin determinanti. Adam Sedgwick, professore di geologia all'Università di Cambridge, fece acquisire a Darwin, nell'estate del 1831, esperienza in osservazioni geologiche sul campo. John Stevens Henslow, professore di botanica a Cambridge e geologo, aggiornò Darwin sugli sviluppi nel settore.

Darwin realizzò la sua seconda grandiosa ambizione in geologia seguendo un percorso più tortuoso; ma, giunto a quel punto, le sue ricerche lo portarono alla pubblicazione più famosa: *On the Origin of Species*. Egli non aveva intrapreso il viaggio sulla *Beagle* con l'intento di scrivere su quell'argomento. Né se lo aspettavano i suoi pari in campo scientifico.

In retrospettiva, si possono identificare nella sua esperienza culturale certe circostanze che lo prepararono a perseguire tale obiettivo, in particolare la tradizione di libero pensiero esistente nella sua famiglia e l'esempio del nonno, Erasmus Darwin, che aveva scritto una cosmologia dello sviluppo. Negli anni trenta del secolo scorso fare propria un'ipotesi evoluzionistica (secondo la quale nuove specie derivano da altre specie preesistenti) voleva comunque dire per Darwin rompere con la maggioranza dei suoi pari: la sua nuova ambizione esigeva una rottura con la tradizione.

Ironia della sorte, ancora una volta fu Lyell a esercitare la maggiore influenza su Darwin. Lyell aveva dedicato la maggior parte del secondo volume di *Principles of Geology* ad attaccare l'ipotesi evoluzionistica dello studioso francese Jean Baptiste de Lamarck. Mentre però l'attaccava, Lyell dava ai suoi lettori una chiara enunciazione dell'ipotesi in questione. Così, quando nel novembre del 1832, mentre la *Beagle* era a Montevideo, ricevette il volume, Darwin ebbe sottomano una presentazione aggiornata e ben informata di entrambi i lati della controversia. Perciò, agli inizi del viaggio egli sapeva qual era l'interpretazione tradizionale riguardante le specie e ciò che avrebbe voluto dire rompere con essa. Inoltre, parlando contro l'evoluzionismo, Lyell aveva presentato una quantità enorme di informazioni sulla distribuzione geografica delle specie e sull'economia della natura, ovvero su quella che oggi verrebbe chiamata ecologia. A queste informazioni Darwin avrebbe attinto più avanti per i propri scopi.

Il problema fondamentale che aveva portato in primo piano in campo geologico la questione dell'evoluzione era quello dell'estinzione. Nel 1796 il naturalista francese Georges Cuvier aveva dichiarato che le ossa di un grande fossile rinvenute in Paraguay appartenevano a un animale sconosciuto, per il quale aveva coniato il nome *Megatherium* (grande belva). Poiché ai suoi tempi i continenti erano stati esplorati abbastanza bene e nessun animale di questo tipo era mai stato avvistato, Cuvier concluse che l'animale si era estinto. Di lì a poco furono identificati i resti di altri animali di grandi dimensioni (*Mastodon* e *Palaeotherium*) senza esemplari viventi. Nel primo Ottocento la tesi dell'estinzione era così convincente da essere considerata una realtà di fatto.

Darwin affrontò il problema dell'estinzione sia come raccoglitore di esemplari sia come interprete. Come raccoglitore si diede da fare durante il viaggio per radunare resti di grandi animali terrestri estinti. Alla fine si ritrovò con resti parziali di una decina circa di forme differenti.

Come interprete, la sua principale preoccupazione per quanto riguardava i fossili fu, durante il viaggio, quella di determinare la loro associazione geologica, vale a dire la loro posizione in uno strato particolare e il rapporto esistente fra quello strato e gli strati vicini. Il suo scopo era quello di stabilire le condizioni in cui le ossa si erano depositate e da ciò acquisire una certa conoscenza delle cause dell'estinzione. Egli si rendeva conto perfettamente della spaccatura esistente nella geologia inglese tra chi (compreso Lyell) credeva che le specie si estinguessero gradualmente e chi (fra cui William Buckland, lettore di geologia all'Università di Oxford) riteneva che le specie scomparissero in modo abbastanza repentino, presumibilmente a causa di mutamenti drammatici delle condizioni di vita dovuti, poniamo, a un'inondazione o a un improvviso cambiamento di temperatura.

Darwin optò per l'idea gradualistica anche per via della gradualità degli innalzamenti della superficie che, a suo parere, stavano avendo luogo nell'America Meridionale. Come scrisse nel 1835 mentre era in Cile: «Per quel che riguarda la *morte* di alcune specie di mammiferi terrestri nella parte sud dell'America Meridionale, sono vivamente propenso a respingere l'azione di qualsiasi calamità improvvisa...».

Poi, come Lyell, proseguiva affermando che le specie estinte dovevano essere state sostituite da nuove specie. Può darsi che questo suo giudizio sorprenda il lettore dei nostri giorni, in quanto si rifaceva alla teologia. «L'idoneità riproduttiva stabilita dall'Autore della natura» - sosteneva - non avrebbe senso se le specie estinte non venissero sostituite. In quell'anno, egli era convinto di avere ragioni teologiche per considerare il modo in cui si sarebbero potute presentare le sostituzioni.

Quando, nell'ottobre del 1836, ritornò in Inghilterra, Darwin portò le ossa fossili che aveva raccolto a Richard Owen, anatomista al Royal College of Surgeons e discepolo di Cuvier. Alla fine di gennaio, Owen aveva già terminato di esaminarle: alcune, annunciò, appartenevano a specie note di animali estinti, mentre altre erano nuove. Ciò che colpì Darwin furono soprattutto le connessioni messe in rilievo da Owen tra le ossa fossili e vari animali viventi tipici dell'America Meridionale. Un esemplare, chiamato in seguito *Toxodon platensis*, aveva affinità con l'attuale capibara. Un altro, chiamato poi *Macrauchenia patachonica*, presentava affinità (alla fine contestate) con l'attuale guanaco. Lo stesso Lyell, poco dopo che Owen si era pronunciato sugli esemplari di Darwin, richiamò l'attenzione su questo impressionante schema di successione geologica nelle forme di animali sudamericani.

A un lettore dei nostri giorni potrà sembrare che, per uno schema di questo genere, non possa esservi che una spiegazione evoluzionistica; nel 1837, a questa si frapponeva, invece, una difficoltà insormontabile. Per quanto numerose fossero le analogie tra forme antiche e forme moderne, vi erano ancora molte differenze ed esistevano varie linee lungo le quali si sarebbero

La teoria di Darwin sulla formazione delle scogliere coralline si basava sul concetto che i coralli crescono su una superficie che sprofonda lentamente. Qui i coralli (*in colore*) si accumulano in acque poco profonde su una superficie che un tempo costituiva il perimetro, sopra il livello dell'acqua, di un'isola che si abbassava (*a*). Alla fine tutta l'isola sprofonda e i coralli formano un atollo: una scogliera intorno a una laguna che giace sopra quella che era la parte più elevata dell'isola (*b*). Darwin spiegò correttamente che i coralli non crescono in una laguna perché per svilupparsi bene hanno bisogno di acqua in movimento, mentre quella lagunare è troppo stagnante.

Le prime due pagine del taccuino sul quale Darwin annotava i campioni geologici che raccoglieva danno un'idea della sua meticolosità nel registrare i dati. Quail Island (oggi Ilhéu de Santa Maria) nelle isole del Capo Verde è il luogo in cui Darwin iniziò a raccogliere campioni geologici e in cui, come scrisse nella sua autobiografia, «mi venne in mente per la prima volta che forse avrei potuto scrivere un libro sulla geologia dei vari paesi visitati». Egli numerava meticolosamente ogni campione e riportava lo stesso numero accanto ai propri commenti su quel campione.

potute dedurre affinità. Non si conosceva, inoltre, alcuna serie di forme intermedie che collegassero le forme antiche con quelle moderne. Perciò, per quanto indicative, le testimonianze fossili non erano sufficienti a sostenere un'ipotesi evoluzionistica. Era necessario un altro elemento.

Questo elemento fu fornito da dati tratti dall'attuale distribuzione geografica delle specie. Il vantaggio delle forme viventi era che presentavano gradazioni impercettibili. Nella sua autobiografia, Darwin, raccontando di come avesse adottato un'ipotesi evoluzionistica, scrisse di essere rimasto colpito sia «dal modo in cui animali strettamente affini si sostituiscono via via che sul continente si procede verso sud», sia «dal carattere sudamericano della maggior parte dell'arcipelago delle Galápagos, e più particolarmente dal modo in cui, in ogni singola isola del gruppo, gli animali presentano lievi differenze ».

Uno degli esempi fondamentali della sostituzione di una specie da parte di un'altra strettamente affine vedeva coinvolte le due specie di nandù: il nandù comune o maggiore (*Rhea americana*), che si trova dal Brasile nord-orientale al Rio Negro nell'Argentina centrale, e il nandù di Darwin o minore (*Rhea darwini*), che si trova nei bassopiani (*pampas*) della Patagonia. L'esempio chiave, portato da Darwin, sulle diversità morfologiche riscontrate da un'isola all'altra era quello dei tordi beffeggiatori delle Galápagos. Tre isole, Charles, Albemarle e Chatham, avevano ognuna un proprio tordo beffeggiatore, a malapena distinguibili tra loro a prima vista, ma che nondimeno, secondo l'ornitologo John Gould, appartenevano a specie diverse.

Nandù, tordi beffeggiatori ed esempi analoghi fecero apparire possibile a Darwin l'evoluzione. Era certo più facile immaginare una forma vivente che si trasformava in un'altra forma vivente leggermente diversa che non immaginare il passaggio da un animale estinto a uno vivente affine, ma nettamente diverso. Una volta capito il potenziale evolutivo delle specie viventi, Darwin lo interpretò rapidamente in chiave del passato. Come annotò sul suo «Taccuino rosso» nel marzo del 1837: «Lo stesso tipo di rapporto che c'è tra lo struzzo maggiore o comune [in realtà il nandù comune] e [il minore]: tra il guanaco estinto e quello recente: nel primo caso il rapporto è dato dalla posizione, nel secondo dal tempo.» Sul taccuino Darwin continuava a considerare la possibilità che una specie si trasformasse in un'altra. Era diventato un evoluzionista.

Così la seconda grande ambizione di Darwin in geologia si cristallizzò nel marzo del 1837, con l'analogia che egli fece paragonando la distribuzione delle specie nello spazio con la loro distribuzione nel tempo. Questa stessa ambizione si realizzò pienamente nel 1859, quando fu pubblicato il libro *On the Origin of Species*. Nei 22 anni intermedi Darwin rimase fedele alla propria vocazione di geologo, ma si dedicò soprattutto alla scrittura, fornendo un resoconto completo sul materiale per il quale era già stato effettuato il lavoro sul campo. Sia come osservatore diretto della natura sia come scrittore il suo retaggio di geologo era comunque assicurato.

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Ramificazioni filogenetiche nel Paleozoico e fortune dei cognomi inglesi

Nell'Ontario del Sud e nella zona occidentale dello stato di New York si trovano distese di rocce del medio Devoniano tra le più belle e le più ricche di fossili del mondo. Una è proprio vicino a casa mia. Lì, in una cava in cui abbondano le tracce di antichi organismi, un giorno mi imbattei in una trilobite che mi fissava da un argilloscisto. Mentre la rimuovevo delicatamente, sentii un rumore e, giratomi di scatto, mi accorsi che alle mie spalle c'era un tipo con la barba che guardava con attenzione quello che stavo facendo. «Un meraviglioso esemplare di *Phacops rana*», sentenziò. «È un vero peccato che siano tutti estinti». Era un professore di paleontologia e si chiamava Smith. «Perché - mi chiesi ad alta voce - le trilobiti si sono estinte?» Il Professor Smith fissò un attimo nel vuoto e rispose alla mia domanda con un'altra domanda: «Perché il cognome Smith è così comune?».

Per rispondere davvero a queste due domande bisognerebbe risalire a un complesso e irreversibile intreccio di fattori genetici e ambientali. Una risposta approssimativa, però, ce la possono dare due programmi per calcolatore che simulano in modo analogo l'evoluzione delle trilobiti e dei cognomi. Uno dei due programmi cerca di simulare le grandi estinzioni del Paleozoico. Venne ideato in origine da David M. Raup, noto paleontologo dell'Università di Chicago, e ripercorre in modo casuale lo sviluppo di un albero filogenetico: a ogni tic dell'orologio geologico, l'estremità di ogni ramo cresce, si ramifica o muore secondo una certa probabilità. La configurazione evolutiva casuale generata dal programma ha una sorprendente somiglianza con la documentazione fossile ed entrambe mostrano estinzioni in massa.

Il secondo programma imita la proliferazione e l'estinzione di cognomi inglesi dal 1350 ai giorni nostri. Il programma è frutto della mente di Christopher M. Sturges e Brian C. Haggett, entrambi funzionari del Ministero della Difesa inglese. Nel loro tempo libero, Sturges e Haggett si occupano di genealogie; essi erano rimasti perplessi dalla graduale scomparsa di molti cognomi dalle documentazioni che avevano potuto consultare. Era possibile che questa scomparsa fosse dovuta al caso? I due pensarono di risolvere il problema scrivendo un programma di simulazione e scoprirono che i tre quarti dei cognomi che erano comuni nel 1350 scomparivano per effetto del solo caso nei successivi 636 anni.

I lettori che vogliano viaggiare in una macchina del tempo probabilistica possono scrivere una versione semplificata dei due programmi; questa riesce comunque a riprodurre le caratteristiche essenziali della versione originale. Nel primo programma passerete attraverso centinaia di milioni di anni osservando generi o interi phyla diversificarsi o morire. Nel secondo programma trascorrerà solo qualche centinaio di anni, ma l'interesse è più umano. Quali cognomi trionferanno, gli Smith o gli Smedley?

Il programma che io chiamo PALEOTREE inizia con un'unica creatura primordiale. Per conservare il sapore paleontologico della simulazione, mi sembra giusto dare alla creatura un nome scientifico: *Paleoplasmus radiculus*. I lettori sono liberi di abbandonarsi ad analoghi abusi del latino. Per un arbitrario tempo dato, PALEOTREE rappresenta la popolazione totale di *P. radiculus* come un unico nodo, o punto, su un albero. Sempre come un nodo viene rappresentata ogni popolazione di creature appartenenti alla stessa categoria biologica di *P. radiculus* e da essa discendente: il programma non fa alcun tentativo di contare il numero di esemplari viventi della categoria biologica rappresentati dal nodo in un determinato momento.

In un istante, si fanno passare milioni di anni e il programma esamina i risultati dell'evoluzione. Può darsi che la vecchia popolazione mantenga la sua integrità biologica per qualche altro milione di anni. Se così è, il nodo originario dà luogo a un nuovo nodo, collegato al suo progenitore da una linea. Può darsi, invece, che la popolazione originaria abbia successo e si diversifichi in una nuova categoria biologica. Questo esito può essere rappresentato da una ramificazione a forma di Y all'altezza del nodo originario. Un ramo della Y porta a un nuovo nodo che rappresenta la continuazione della categoria biologica originale e il secondo ramo porta a un altro nodo che rappresenta l'inizio di una nuova categoria. L'ultima possibilità è che la popolazione originaria si estingua. In questo caso il nodo originario termina e da esso non si diparte alcun ramo.

PALEOTREE determina a caso il destino di ogni nodo che genera nell'albero: la popolazione rimarrà immutata, darà luogo a una nuova categoria biologica oppure si estinguerà? Prese così come sono, le tre possibilità non riflettono realmente la ricca struttura dell'evoluzione biologica. Così come ho descritto il programma finora, tutte le creature discendenti da un comune antenato sono monofiletiche: in altri termini, sono in rapporti abbastanza stretti da poterle includere nella stessa ampia categoria biologica e quindi appartengono allo stesso albero. Lo sviluppo evolutivo, però, porta di tanto in tanto a una nuova ampia categoria o a una estinzione in massa quando intere categorie scompaiono. I tassonomisti riconoscono molte categorie di questo genere, dalle varietà e sottospecie ai phyla e ai regni. Per semplicità PALEOTREE riconosce solo due categorie, la specie e il genere; andrebbe altrettanto bene qualsiasi altra interpretazione di queste due categorie.

In PALEOTREE un genere può dare origine a un intero nuovo genere. Il meccanismo è semplice: una nuova specie si ramifica dal vecchio genere per fondarne uno nuovo. Quello che, in realtà, è solo un altro ramo dell'albero evolutivo è visto come un albero a sé stante. Naturalmente, un albero (un genere) può anche estinguersi: tutte le sue specie muoiono. Il numero di specie che formano un genere a un dato momento è la diversità di quel genere e misura non solo la ricchezza genetica del genere, ma anche la sua probabilità di sopravvivenza. C'è un metodo elegante per rappresentare graficamente i risultati di un esperimento con PALEOTREE: ogni genere viene rappresentato da una pila simmetrica di sbarrette orizzontali che ricorda un fuso. Ogni sbarretta rappresenta una nuova epoca e la sua lunghezza a un dato livello della pila corrisponde alla diversità del genere nell'epoca corrispondente. La pila risultante mi ricorda un bizzarro sistema di illuminazione domestica (*si veda l'illustrazione di queste due pagine*).

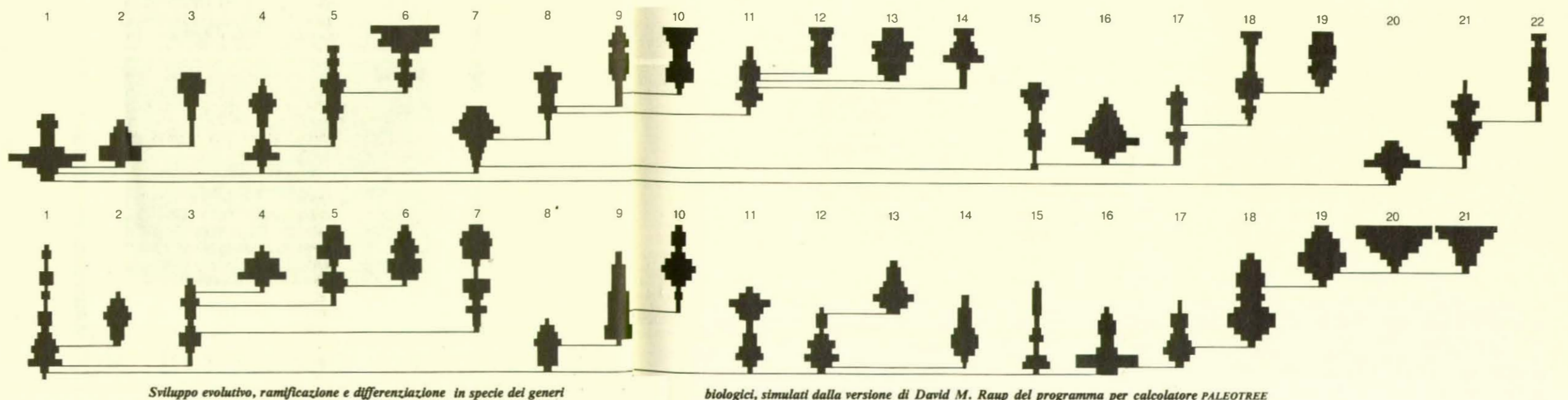
È facile e istruttivo generare a mano un albero e il fuso a esso associato. Si tracci un punto su un pezzo di carta e si lanci una coppia di dadi. Se la somma dei numeri che compaiono sui dadi è 2, 3 o 4, la specie si estingue immediatamente: *P. radiculus* è nato morto e nessun ramo emerge dal punto. Se la somma è 9 o 10, la specie si diversifica: dal punto si dipartono verso l'alto due corte linee e si segna un punto all'estremità di ciascuna linea. Se la somma è pari a qualsiasi altro numero, la specie non cambia: un'unica linea sale verso l'alto e termina con un altro punto. Per ogni epoca, si ripete l'esperimento per ogni specie che non sia ancora estinta. Il procedimento dà origine a un albero e il numero di punti a un dato livello dell'albero può essere rappresentato come la lunghezza di una sbarretta nel diagramma a fuso (*si veda l'illustrazione della pagina successiva*).

Quando si lanciano i dadi, la probabilità che una specie si estingua nell'epoca successiva è 1/6 (ossia 0,167), la probabilità che si diversifichi è 7/36 (ossia 0,194) e la probabilità che rimanga immutata è di 23/36 (ossia 0,639). Queste probabilità sono vicine a quelle ipotizzate in PALEOTREE e danno luogo ad alberi analoghi. Dato che le probabilità di estinzione sono un po' inferiori a quelle di diversificazione, ci si potrebbe aspettare che qualche albero proceda all'infinito. È davvero così? Sarei lieto di sapere che qualche lettore è rimasto alzato tutta la notte per verificarlo.

Per disegnare un diagramma a fuso non è necessario conservare un albero per ogni genere. PALEOTREE si limita a contare il numero di specie che costituiscono un genere in ogni epoca e conserva una registrazione delle ramificazioni tra i generi. Entrambe le funzioni possono essere controllate da una doppia matrice chiamata *storia*. Le colonne rappresentano i generi e le righe le epoche. *Storia*(i,j), quindi, è il numero di specie presenti al momento i nel j-esimo genere. Una volta che PALEOTREE ha simulato l'evoluzione per il numero di epoche desiderato, si può stampare la matrice *storia* e usare i suoi elementi per disegnare a mano i diagrammi a fuso. I lettori che abbiano qualche esperienza di programmazione grafica potrebbero tentare di aggiungere un programma che disegni automaticamente il diagramma a fuso, ma una descrizione di questo esercizio richiederebbe un intero articolo.

Per far girare PALEOTREE nella sua forma più semplice si deve battere, all'inizio del programma, il numero di epoche, o «generazioni». Tale numero viene conservato dal programma sotto il nome *numgen* e limita il valore dell'indice i per la matrice *storia*. Mentre il programma esegue il suo ciclo esterno, il valore di i aumenta con incrementi unitari da 1 a *numgen*. In un secondo ciclo, interno al primo, si utilizza come indice il valore di j della matrice *storia*. Ogni valore di j corrisponde a un genere. Si definisce anche una variabile *max* per calcolare il limite dell'indice j. Dato che all'inizio è presente solo un genere, *max* inizialmente è posto uguale a 1. Eseguendo ripetutamente il ciclo esterno, il numero di generi aumenta e diminuisce. Per ogni nuova epoca, si pone *max* uguale al numero di generi dell'epoca precedente. Una variabile detta *lmax* è posta uguale al valore di *max* all'inizio del ciclo j e j va da 1 a *lmax*.

Nel secondo ciclo, PALEOTREE prende in esame il numero di specie attualmente incluse nel genere j-esimo e lo immagazzina in una variabile temporanea detta *temp*. Il programma entra poi nel terzo ciclo, quello più interno, che stabilisce il destino di ogni specie del genere j-esimo. L'indice k del ciclo interno può essere considerato un contatore per la specie: va da 1 a *temp* e il valore di *temp* viene trasferito anche all'elemento *storia*$(i+1,j)$ della matrice, che è destinato a fornire il numero di specie del j-esimo genere nella successiva epoca $i+1$.

Per ogni specie, o in altre parole per ogni valore di k, si sceglie a caso un numero compreso tra 0 e 1 per mezzo dell'istruzione del proprio linguaggio di pro-

Sviluppo evolutivo, ramificazione e differenziazione in specie dei generi biologici, simulati dalla versione di David M. Raup del programma per calcolatore PALEOTREE

grammazione per la generazione di numeri a caso. Se il numero casuale è minore di 0,15, la k-esima specie si estingue e $storia(i+1, j)$ diminuisce di un'unità. Se il numero scelto a caso è maggiore di 0,8, la k-esima specie produce un ramo evolutivo e $storia(i+1, j)$ viene aumentata di un'unità. Il procedimento assicura che la probabilità di estinzione di una specie è pari a 0,15 e quella di generazione di una nuova specie è pari a 0,2.

Il programma scritto da Raup, su cui si basa PALEOTREE, determinava nuovi generi attraverso una forma di analisi retrospettiva che sarebbe un poco troppo complicato descrivere in questa sede. Raup ha notato, però, che si possono ottenere risultati ragionevolmente simili in modo più semplice: ammettiamo che ogni specie dia origine a un nuovo genere con la probabilità piuttosto bassa di 0,2 e incorporiamo questo presupposto nel ciclo più interno. Se il numero casuale che controlla il destino della k-esima specie è maggiore di 0,98, PALEOTREE compie una momentanea deviazione per creare un nuovo genere, anziché una nuova specie.

Numerosi dettagli servono a rendere più economico il programma in questa fase. In primo luogo max si aumenta di 1 e l'i-esimo elemento della nuova colonna della matrice *storia*, cioè $storia(i, max)$, è posto uguale all'indice j del genere che lo ha originato. In questo modo la matrice può immagazzinare, all'inizio del genere appena creato, informazioni relative all'identità del genere che l'ha originato. Il lettore (o un programma per la grafica) può quindi stabilire l'origine di ogni genere.

Nell'epoca $i + 1$, il primo stadio della sua reale evoluzione, il genere appena creato ha solo una specie. Quindi l'elemento $storia(i+1, max)$ è posto uguale a 1 e da qui l'evoluzione procede normalmente.

Date le probabilità stabilite nel programma, quasi inevitabilmente alcuni generi si estinguono. In questo caso, la colonna di *storia* corrispondente al genere che si è appena estinto diventa 0 e max deve essere diminuito di 1. Tale circostanza può creare difficoltà in alcuni linguaggi di programmazione. Anziché ignorare effettivamente la colonna, il ciclo più interno può venir eseguito una o due volte anche se *temp* è uguale a 0. Alcuni linguaggi possono presupporre che questo ciclo proceda in modo decrescente, da 1 a 0, e quindi può accadere che il genere abbia una o due occasioni per una rinascita miracolosa. Per evitare questa situazione, si può inserire una verifica nella parte iniziale del ciclo più interno: se *temp* è uguale a 0, il resto del ciclo va saltato.

Come sempre avviene, i lettori avranno forse voglia di mettere le mani nei programmi proposti in questa rubrica. Per esempio, che cosa accade alla diversità se si aumenta da 0,15 a 0,21 la probabilità di estinzione? Il fatto che la probabilità di ramificazione sia uguale a quella di estinzione porterà all'estinzione totale della vita?

Non dovrebbe sorprendere nessuno il fatto che il secondo programma, che io chiamo NEOTREE, sia del tutto simile a PALEOTREE nella struttura e nella funzione. Si ricordi che esso riproduce l'estinzione e la sopravvivenza di cognomi. In effetti lo si può ottenere da PALEOTREE eliminando la procedura per la nascita di nuovi generi.

Anche in NEOTREE una matrice chiamata *storia* segue gli eventi al passare del tempo. Quando il programma parte, però, *storia* contiene 1000 elementi anziché uno solo. Gli elementi rappresentano 1000 cognomi inglesi molto diffusi nel 1350. Dato che sarebbe folle (anche se non privo di un certo fascino d'antiquariato) incorporare 1000 cognomi veri nel programma, ogni colonna di *storia* rappresenta un'unica famiglia il cui cognome è dato, purtroppo, dall'indice numerico della colonna. L'elemento iniziale di ciascuna delle 1000 colonne è 1, che rappresenta un nucleo familiare inglese: padre, madre e numerosi figli. Di fatto, il programma segue solo i maschi, cioè i membri della famiglia che trasmettono il cognome alla generazione successiva.

Sarebbe ben poco realistico imitare il programma PALEOTREE e limitare una certa linea di discendenza a due nuove famiglie per ogni generazione. In NEOTREE, quindi, a ogni famiglia è consentito generare da zero a sei nuove famiglie. Il ciclo più interno di NEOTREE opera sui cognomi allo stesso modo in cui il ciclo più interno di PALEOTREE opera sui generi. Quando NEOTREE considera il cognome j-esimo, per esempio, la variabile *temp* è posta uguale al numero di famiglie che in quel momento hanno quel cognome. Il ciclo più interno considera volta a volta ogni famiglia e un numero casuale stabilisce poi il numero di maschi di quella famiglia che si sposano.

Sturges e Hagget impiegano una tabella derivata da un'analisi statistica dei documenti genealogici per stabilire le probabilità di sette possibili destini di cognomi (*si veda l'illustrazione nella pagina a fronte*). Per esempio, si ha una probabilità pari a 0,317 che una famiglia non produca maschi destinati a sposarsi, mentre è pari a 0,364 la probabilità che essa produca un solo maschio che si sposa, e così via.

È facile trasformare i numeri casuali generati dal programma in una tabella in cui vengano riprodotte le probabilità che una famiglia generi un certo numero di maschi che si sposano. Per esempio, se il numero casuale è compreso tra 0 e 0,317, la famiglia non avrà maschi che si sposano ed essa (se non il cognome stesso) si «estinguerà». Se il numero casuale è compreso tra 0,317 e 0,681 (la somma di 0,317 e 0,364), la famiglia genererà un maschio che si sposa. In altre parole, il numero casuale apparterrà al secondo campo numerico circa 364 volte su 1000, e questa proporzione è in accordo con la probabilità che si ricava dall'analisi statistica. Il resto della tabella è costruito in modo analogo.

Una volta che ha stabilito il destino di una certa famiglia per una nuova generazione, NEOTREE si comporta esattamente come PALEOTREE. All'inizio dell'i-esima generazione, il numero di famiglie con cognome j è immagazzinato nella variabile *temp*. L'indice k del ciclo più interno va poi da 1 a *temp* e *temp* è immagazzinato in $storia(i+1, j)$, proprio come avviene in PALEOTREE. Per ogni famiglia di cognome j che non produce maschi che si sposano, $storia(i+1, j)$ diminuisce di un'unità, e questo riflette il fatto che una famiglia in meno porterà il cognome j nella generazione successiva. Per ogni famiglia di cognome j che produce un maschio che si sposa, $storia(i+1, j)$ rimane immutata. Per ogni famiglia di cognome j con due maschi che si sposano, il valore di $storia(i+1, j)$ è aumentato di un'unità, e così via.

In un certo senso, NEOTREE e PALEOTREE portano a risultati opposti. In NEOTREE, per esempio, si consideri ogni generazione pari a 28 anni; il periodo che va dal 1350 al 1986 richiede quindi 22 o 23 generazioni di calcolo. Quanti cognomi scompaiono in questo periodo? Sturges e Haggett hanno scoperto che il numero dei cognomi diminuisce costantemente e che quasi tre quarti scompaiono nell'arco dei 636 anni. Forse qualche lettore avrà la tenacia di far girare NEOTREE tanto da prevedere l'anno in cui i 1000 cognomi originari saranno ridotti a uno solo!

I risultati di PALEOTREE sono ben diversi: come era prevedibile, e nonostante le estinzioni di generi, la vita continua a diversificarsi e con il passare del tempo il numero di generi tende a crescere. Ci si potrebbe aspettare che alcuni si arricchiscano talmente di specie da risultare in pratica immuni da ogni rischio di estinzione prima che l'universo collassi. Eppure, nella versione iniziale di Raup si aveva un numero inconsueto di estinzioni: un esito davvero sconcertante.

PALEOTREE è stato sviluppato negli anni settanta, quando un gran numero di scienziati hanno rivolto l'attenzione alle chiare prove di estinzioni in massa fornite dalla documentazione fossile. Se si poteva prestare fede alle prove, le estinzioni in massa sarebbero state causate da catastrofi geologiche o astrofisiche. Lavorando con Jay Gould dell'Università di Harvard e con altri, Raup ha avanzato l'ipotesi che, con un programma di simulazione veloce come PALEOTREE, si potesse procedere per esclusione: se esso avesse generato estinzioni in massa simili a quelle osservate, non sarebbe più stata così pressante la necessità di spiegarle come effetti di una catastrofe.

Per giustizia nei confronti di Raup, devo rilevare che egli, prima di scrivere il programma, non aveva alcuna ragione per dubitare dell'esistenza di estinzioni in massa. Raup si aspettava di trovare scarsa o nulla correlazione tra i risultati del suo programma e le ampie fluttuazioni delle testimonianze fossili.

In realtà, invece, ha scoperto che la correlazione era notevolmente elevata. Alcune linee filogenetiche generate dal programma si ramificavano e si ampliavano, mentre altre si esaurivano, proprio come avviene nelle testimonianze fossili. Si presentavano anche casi di estinzione contemporanea di parecchie linee. Che bisogno c'era di appellarsi a un'epoca glaciale o alla collisione di un asteroide se il caso bastava a spiegare le testimonianze di repentine scomparse in natura?

Poco dopo la pubblicazione dei risultati di Raup, Steven M. Stanley della Johns Hopkins University ha formulato una critica efficace all'interpretazione che Raup aveva fornito di essi, rilevando che il numero di specie ammesso dal programma per un genere era troppo piccolo per essere realistico. Più basso è questo numero, più facilmente le fluttuazioni probabilistiche portano all'estinzione del genere. In un genere vero, il numero delle specie è abbastanza alto da sopportare la fluttuazione. A dispetto, quindi, delle estinzioni in massa presentate dalla simulazione di Raup, quelle testimoniate dai fossili potrebbero non essere un prodotto statistico. Raup ha riconosciuto subito la validità della critica di Stanley e ora è un sostenitore della teoria secondo cui le estinzioni in massa non risultano da variazioni casuali.

Una ragione per la ridotta diversità di specie che si trova nei risultati di Raup era che il programma dichiarava inizialmente matrici piuttosto piccole, che limitavano artificialmente ogni genere a un massimo di 40 specie. Secondo le parole di Raup: «Non ci rendemmo conto a quel tempo di quanto fosse importante la dimensione». Il suo programma non inseriva più di una specie in un singolo elemento della matrice, come fa PALEOTREE, perché Raup voleva ricostruire l'intero albero evolutivo per ogni specie e poi analizzare da vicino il processo della speciazione casuale.

Un'altra ragione della ridotta diversità di specie è che sia il programma di Raup sia PALEOTREE iniziano la costruzione evolutiva a partire da un'unica specie. Un modello simile potrebbe avvicinarsi alle primissime fasi della vita, quando la diversità di specie era bassa. I lettori che vogliano vedere in azione la critica di Stanley dovrebbero fornire inizialmente a *storia* 100 generi, ciascuno con 100 specie. Si incontrano ancora estinzioni in massa? Darò notizia di ogni evento inconsueto, sia per le trilobiti sia per gli Smith.

Speciazione all'interno di un genere, determinata dal lancio di una coppia di dadi

NUMERO DI MASCHI CHE SI SPOSANO	0	1	2	3	4	5	6
PROBABILITÀ	0,317	0,364	0,209	0,080	0,023	0,005	0,001

Riassunto statistico, a partire dal 1350, dei maschi inglesi che hanno contratto matrimonio

Nell'articolo del mese scorso ho parlato della follia quadrimensionale di Magi. Il mio assistente algoritmico sosteneva di essere scivolato nella quarta dimensione e di avere scoperto alcune notevoli proprietà degli ipercubi. Quanto è lunga, chiedeva, la diagonale di un ipercubo a n dimensioni? Che cosa accade quando una ipersfera a n dimensioni è iscritta in un ipercubo a n dimensioni? Pensavo che le sue risposte fossero folli perché sembravano andare contro l'intuizione comune. La diagonale si allunga illimitatamente al crescere della dimensione dell'ipercubo. Al contempo, il volume dell'ipersfera iscritta va progressivamente diminuendo!

La prima risposta è facilmente derivabile estendendo la formula euclidea per la distanza, dalle due dimensioni a un numero superiore di dimensioni: la lunghezza della diagonale dell'ipercubo a n dimensioni è semplicemente la radice quadrata di n e quel numero cresce senza limite con l'aumentare di n.

Non conosco un modo elementare per dimostrare la contrazione di un'ipersfera, ma Thomas Banchoff della Brown University, cui rubrica del mese scorso mostrava alcune rappresentazioni grafiche di una ipersfera in rotazione, fornisce un'argomentazione plausibile.

Il volume dell'iperottaedro iscritto è grosso modo paragonabile a quello dell'ipersfera. Ma come si caratterizza l'iperottaedro iscritto? Ogni vertice di un comune ottaedro iscritto in un cubo tocca una faccia del cubo. Analogamente, ogni vertice dell'iperottaedro sta al centro di un'iperfaccia dell'ipercubo. Con questo suggerimento, non è difficile calcolare il volume dell'iperottaedro, volume che diminuisce con il decrescere del numero delle dimensioni.

Scherzi primaverili a parte, non devo a Magi questi affascinanti problemi, bensì al temibile matemago Ronald L. Graham degli AT&T Bell Laboratories.

Nell'articolo di aprile ho descritto la macchina che gioca a scacchi, costruita nel 1890 da Leonardo Torres y Quevedo, un ingegnere spagnolo. La macchina riusciva sempre a dare scacco matto al re nero dell'avversario umano con la sua combinazione di torre e re. Come fece Torres y Quevedo a costruirla? Se ne può trovare una vecchia ma eccellente descrizione in un libro di David Levy, *Chess and Computers*. Come si possono ideare le mosse di una macchina del genere? Io avevo invitato i lettori a proporre algoritmi altrettanto inesorabili.

I modi di affrontare il problema, per non parlare degli stili algoritmici, sono stati così differenti che non ho potuto individuare un'unica soluzione come la più rapida. Tre lettori, Scott K. Liddell di Silver Spring, Maryland, Paul Canniff di Marlton, New Jersey, e Stephen J. Perris di Baton Rouge, Louisiana, non solo hanno usato algoritmi simili per arrivare allo

scacco matto, ma hanno anche scritto programmi per confermarne l'efficacia.

L'algoritmo di Liddell, il migliore dei tre, applica una strategia di rotazioni e divisioni: «Ruota la scacchiera per trovare ogni configurazione in cui il re nero sia almeno due traverse avanti al re bianco. C'è sempre almeno una configurazione del genere, ma per lo più ce ne sono due. Se l'algoritmo trova una sola di queste configurazioni, la scelta è obbligata. Se vengono trovate due configurazioni, il programma sceglie quella in cui il re nero è più vicino all'estremità della scacchiera. Una volta scelta una posizione, il programma taglia la scacchiera in modo che il re isolato non possa avanzare. Se si trova in una posizione che le impedisce di tagliare l'avanzata del re nero, la torre bianca si sposta all'estremità della scacchiera. Qualsiasi cosa faccia il Nero, il programma è poi in grado di tagliare la scacchiera alla mossa successiva. L'intera procedura richiede almeno due mosse da parte del Bianco.»

Con «intera procedura» ci si riferisce a un ciclo base di mosse con le quali il Nero viene obbligato a spostarsi progressivamente verso un angolo della scacchiera. Un grande vantaggio dell'algoritmo sottostante è l'uso della simmetria di rotazione; se il re nero non si trova «al di sopra» del re bianco, il programma non fa che ruotare la scacchiera fino a ottenere la posizione desiderata.

Numerose fra le strategie proposte dai lettori applicano un concetto chiave degli scacchi: l'opposizione. Se i due re distano solo due caselle uno dall'altro, il re che deve muoversi non si può avvicinare all'altro re: deve o mantenere la stessa distanza oppure allontanarsi. Forse la migliore delle strategie che impiegano una continua opposizione è quella proposta da Alexis A. Gilliland di Arlington, Virginia. Ecco le regole (lievemente parafrasate) che adotta.

1. La torre occupa sempre la prima o l'ottava traversa. Quando viene attaccata su una traversa, si sposta sull'altra.

2. Il re bianco si sposta verso il centro della scacchiera e cerca di spingere il re nero verso la colonna della torre.

3. Il Bianco cerca una posizione di opposizione in cui i due re si trovino distanti una colonna sulla stessa traversa e tocchi al Bianco muoversi. Per arrivare all'opposizione può essere necessario che il Bianco effettui una mossa inutile con la torre, violando temporaneamente la regola 1.

4. Una volta ottenuta l'opposizione, il Bianco dà scacco al re nero con la torre spostandola sulla colonna del re bianco e obbligando così il re nero ad avvicinarsi di una colonna al margine della scacchiera.

5. Si ripetono i passi 3 e 4 finché il re nero arriva al margine della scacchiera. A quel punto il passo 4 porta allo scacco matto.

Gilliland sostiene che la sua strategia conduce inevitabilmente allo scacco matto e afferma anche che difficilmente un algoritmo per arrivare allo scacco matto nel minor numero di mosse ha a sua volta una lunghezza minima.